GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2005, 18:04:25 ; Search time 169 Seconds
(without alignments)
2106.453 Million cell updates/sec

Title: US-10-679-102-15
Perfect score: 4524
Sequence: 1 MLGPAVLGLSLWALLHPGTG.....GPGDAQGQNDGNTGNQGKHE 852

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 4524 | 100.0 | 852 | 3 | US-09-897-427A-6 | Sequence 6, Appli |
| 2 | 4524 | 100.0 | 852 | 3 | US-09-927-315-15 | Sequence 15, Appl |
| 3 | 4524 | 100.0 | 852 | 4 | US-10-035-045-4 | Sequence 4, Appli |
| 4 | 4524 | 100.0 | 852 | 4 | US-10-190-417-15 | Sequence 15, Appl |
| 5 | 4524 | 100.0 | 852 | 4 | US-10-725-276-4 | Sequence 4, Appli |
| 6 | 4524 | 100.0 | 852 | 5 | US-10-725-284-4 | Sequence 4, Appli |
| 7 | 4524 | 100.0 | 852 | 5 | US-10-679-102-15 | Sequence 15, Appl |
| 8 | 4524 | 100.0 | 936 | 4 | US-10-292-798-602 | Sequence 602, App |
| 9 | 4517 | 99.8 | 852 | 4 | US-10-188-186-148 | Sequence 148, App |
| 10 | 4512 | 99.7 | 852 | 3 | US-09-796-338A-14 | Sequence 14, Appl |
| 11 | 4512 | 99.7 | 852 | 4 | US-10-261-482-2 | Sequence 2, Appli |
| 12 | 4512 | 99.7 | 852 | 4 | US-10-282-837-14 | Sequence 14, Appl |
| 13 | 4512 | 99.7 | 852 | 4 | US-10-145-586-14 | Sequence 14, Appl |
| 14 | 4512 | 99.7 | 852 | 4 | US-10-407-079-90 | Sequence 90, Appl |
| 15 | 4512 | 99.7 | 852 | 4 | US-10-179-373-7 | Sequence 7, Appli |
| 16 | 4512 | 99.7 | 852 | 4 | US-10-725-103-7 | Sequence 7, Appli |
| 17 | 4512 | 99.7 | 852 | 4 | US-10-725-489-7 | Sequence 7, Appli |
| 18 | 4512 | 99.7 | 852 | 4 | US-10-725-080A-7 | Sequence 7, Appli |
| 19 | 4512 | 99.7 | 852 | 4 | US-10-725-472A-7 | Sequence 7, Appli |
| 20 | 4512 | 99.7 | 852 | 4 | US-10-770-127-199 | Sequence 199, App |
| 21 | 4512 | 99.7 | 852 | 5 | US-10-725-418-7 | Sequence 7, Appli |
| 22 | 4509 | 99.7 | 852 | 4 | US-10-280-183A-5 | Sequence 5, Appli |
| 23 | 4506 | 99.6 | 852 | 5 | US-10-475-620A-7 | Sequence 7, Appli |
| 24 | 4506 | 99.6 | 852 | 5 | US-10-892-632-2 | Sequence 2, Appli |
| 25 | 4500 | 99.5 | 850 | 3 | US-09-799-629-4 | Sequence 4, Appli |
| 26 | 4500 | 99.5 | 850 | 4 | US-10-190-417-30 | Sequence 30, Appl |
| 27 | 4500 | 99.5 | 850 | 4 | US-10-725-081-4 | Sequence 4, Appli |
| 28 | 4500 | 99.5 | 850 | 4 | US-10-724-223-4 | Sequence 4, Appli |
| 29 | 4500 | 99.5 | 850 | 4 | US-10-724-222-4 | Sequence 4, Appli |
| 30 | 4500 | 99.5 | 850 | 5 | US-10-726-568-4 | Sequence 4, Appli |
| 31 | 4500 | 99.5 | 850 | 5 | US-10-679-102-31 | Sequence 31, Appl |
| 32 | 4496.5 | 99.4 | 863 | 4 | US-10-436-715-2 | Sequence 2, Appli |
| 33 | 4496.5 | 99.4 | 863 | 4 | US-10-311-196-2 | Sequence 2, Appli |
| 34 | 4490 | 99.2 | 850 | 4 | US-10-246-785-7 | Sequence 7, Appli |
| 35 | 4474 | 98.9 | 894 | 4 | US-10-436-715-6 | Sequence 6, Appli |
| 36 | 4461 | 98.6 | 854 | 4 | US-10-159-339-2 | Sequence 2, Appli |
| 37 | 4461 | 98.6 | 854 | 4 | US-10-436-715-86 | Sequence 86, Appl |
| 38 | 4365.5 | 96.5 | 829 | 4 | US-10-436-715-4 | Sequence 4, Appli |
| 39 | 4142 | 91.6 | 1577 | 4 | US-10-017-161-690 | Sequence 690, App |
| 40 | 4002 | 88.5 | 771 | 4 | US-10-225-567A-593 | Sequence 593, App |
| 41 | 3708 | 82.0 | 855 | 4 | US-10-297-021-1 | Sequence 1, Appli |
| 42 | 3255 | 71.9 | 858 | 3 | US-09-799-629-14 | Sequence 14, Appl |
| 43 | 3255 | 71.9 | 858 | 3 | US-09-927-315-25 | Sequence 25, Appl |
| 44 | 3255 | 71.9 | 858 | 4 | US-10-035-045-14 | Sequence 14, Appl |
| 45 | 3255 | 71.9 | 858 | 4 | US-10-246-785-9 | Sequence 9, Appli |

ALIGNMENTS

RESULT 1
US-09-897-427A-6
; Sequence 6, Application US/09897427A
; Patent No. US20020160424A1
; GENERAL INFORMATION:
; APPLICANT: ADLER, JON ELLIOT
; APPLICANT: LI, XIADONG
; APPLICANT: STAZEWSKI, LENA
; APPLICANT: XU, HONG
; APPLICANT: EHEVERRI, FERNANDO
; TITLE OF INVENTION: T1R HETERO-OLIGOMERIC TASTE RECEPTORS
; FILE REFERENCE: 078003-0282558
; CURRENT APPLICATION NUMBER: US/09/897,427A
; CURRENT FILING DATE: 2001-07-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 852
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-897-427A-6

Query Match 100.0%; Score 4524; DB 3; Length 852;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 852; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
```

```
Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420

Qy        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540

Qy        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600

Qy        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720

Qy        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780

Qy        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840

Qy        841 NDGNTGNQGKHE 852
              ||||||||||||
Db        841 NDGNTGNQGKHE 852
```

RESULT 2
US-09-927-315-15
; Sequence 15, Application US/09927315
; Publication No. US20030040045A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Ryba, Nicholas J.P.
; APPLICANT: Nelson, Greg
; APPLICANT: Hoon, Mark A.
; APPLICANT: Chandrashekar, Jayaram
; APPLICANT: Zhang, Yifeng
; APPLICANT: The Regents of the University of California
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by the Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: Mammalian Sweet Taste Receptors
; FILE REFERENCE: 02307E-120110US
; CURRENT APPLICATION NUMBER: US/09/927,315
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: US 60/302,898
; PRIOR FILING DATE: 2001-07-03
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 852
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human T1R3 sweet taste receptor
US-09-927-315-15

Query Match 100.0%; Score 4524; DB 3; Length 852;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 852; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360

Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420

Qy        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540

Qy        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600

Qy        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720

Qy        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780

Qy        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840

Qy        841 NDGNTGNQGKHE 852
              ||||||||||||
Db        841 NDGNTGNQGKHE 852
```

RESULT 3
US-10-035-045-4
; Sequence 4, Application US/10035045
; Publication No. US20030054448A1
; GENERAL INFORMATION:
; APPLICANT: ADLER, JON ELLIOT
; APPLICANT: LI, XIAODONG
; APPLICANT: STASZEWSKI, LENA
; APPLICANT: O'CONNELL, SHAWN
; APPLICANT: ZOZULYA, SERGEY
; TITLE OF INVENTION: T1R TASTE RECEPTORS AND GENES ENCODING SAME
; FILE REFERENCE: 078003-0280681
; CURRENT APPLICATION NUMBER: US/10/035,045

```
;  CURRENT FILING DATE:  2002-01-03
;  PRIOR APPLICATION NUMBER: 60/259,227
;  PRIOR FILING DATE: 2001-01-03
;  PRIOR APPLICATION NUMBER: 60/284,547
;  PRIOR FILING DATE: 2001-04-19
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 852
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-035-045-4

  Query Match             100.0%;  Score 4524;  DB 4;  Length 852;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  852;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360

Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420

Qy        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540

Qy        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600

Qy        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720

Qy        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780

Qy        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840

Qy        841 NDGNTGNQGKHE 852
              ||||||||||||
Db        841 NDGNTGNQGKHE 852


RESULT 4
US-10-190-417-15
; Sequence 15, Application US/10190417
; Publication No. US20030166137A1
; GENERAL INFORMATION:
;  APPLICANT: Zuker, Charles S.
;  APPLICANT:  Ryba, Nicholas J. P.
;  APPLICANT:  Chandrashekar, Jayaram
;  APPLICANT:  Hoon, Mark A.
;  APPLICANT:  Nelson, Greg
;  APPLICANT:  Zhang, Yifeng
;  APPLICANT:  The Regents of the University of California
;  APPLICANT:  The Government of the United States of America
;  APPLICANT:  as represented by the Secretary of the
;  APPLICANT:  Department of Health and Human Services
;  TITLE OF INVENTION: Mammalian Sweet and Amino Acid Heterodimeric Taste
;  TITLE OF INVENTION:  Receptors
;  FILE REFERENCE: 02307E-120130US
;  CURRENT APPLICATION NUMBER: US/10/190,417
;  CURRENT FILING DATE:  2002-11-14
;  PRIOR APPLICATION NUMBER: US 60/302,898
;  PRIOR FILING DATE: 2001-07-03
;  PRIOR APPLICATION NUMBER: US 09/927,315
;  PRIOR FILING DATE: 2001-08-10
;  PRIOR APPLICATION NUMBER: US 60/358,925
;  PRIOR FILING DATE: 2002-02-22
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
;   LENGTH: 852
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: human T1R3
US-10-190-417-15

  Query Match             100.0%;  Score 4524;  DB 4;  Length 852;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  852;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360

Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
```

```
Qy        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540

Qy        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600

Qy        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720

Qy        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780

Qy        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840

Qy        841 NDGNTGNQGKHE 852
              ||||||||||||
Db        841 NDGNTGNQGKHE 852
```

RESULT 5
US-10-725-276-4
; Sequence 4, Application US/10725276
; Publication No. US20040209286A1
; GENERAL INFORMATION:
; APPLICANT: ADLER, JON ELLIOT
; APPLICANT: LI, XIAODONG
; APPLICANT: STASZEWSKI, LENA
; APPLICANT: O'CONNELL, SHAWN
; APPLICANT: ZOZULYA, SERGEY
; TITLE OF INVENTION: T1R TASTE RECEPTORS AND GENES ENCODING SAME
; FILE REFERENCE: 078003-0280681
; CURRENT APPLICATION NUMBER: US/10/725,276
; CURRENT FILING DATE: 2003-12-02
; PRIOR APPLICATION NUMBER: 60/259,227
; PRIOR FILING DATE: 2001-01-03
; PRIOR APPLICATION NUMBER: 60/284,547
; PRIOR FILING DATE: 2001-04-19
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 852
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-725-276-4

Query Match 100.0%; Score 4524; DB 4; Length 852;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 852; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360

Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420

Qy        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540

Qy        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600

Qy        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720

Qy        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780

Qy        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840

Qy        841 NDGNTGNQGKHE 852
              ||||||||||||
Db        841 NDGNTGNQGKHE 852
```

RESULT 6
US-10-725-284-4
; Sequence 4, Application US/10725284
; Publication No. US20050032158A1
; GENERAL INFORMATION:
; APPLICANT: ADLER, JON ELLIOT
; APPLICANT: LI, XIAODONG
; APPLICANT: STASZEWSKI, LENA
; APPLICANT: O'CONNELL, SHAWN
; APPLICANT: ZOZULYA, SERGEY
; TITLE OF INVENTION: T1R TASTE RECEPTORS AND GENES ENCODING SAME
; FILE REFERENCE: 078003-0280681
; CURRENT APPLICATION NUMBER: US/10/725,284
; CURRENT FILING DATE: 2003-12-02
; PRIOR APPLICATION NUMBER: 60/259,227
; PRIOR FILING DATE: 2001-01-03
; PRIOR APPLICATION NUMBER: 60/284,547
; PRIOR FILING DATE: 2001-04-19
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 852

```
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-725-284-4

  Query Match           100.0%;  Score 4524;  DB 5;  Length 852;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  852;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360

Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420

Qy        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540

Qy        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600

Qy        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720

Qy        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780

Qy        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840

Qy        841 NDGNTGNQGKHE 852
              ||||||||||||
Db        841 NDGNTGNQGKHE 852


RESULT 7
US-10-679-102-15
; Sequence 15, Application US/10679102
; Publication No. US20050106571A1
; GENERAL INFORMATION:
;  APPLICANT: Erlenbach, Isolde
;  APPLICANT:  Ryba, Nicholas J. P.
;  APPLICANT:  Zhao, Grace
;  APPLICANT:  Zuker, Charles S.
;  APPLICANT:  The Regents of the University of California
;  APPLICANT:  The Government of the United States of America
;  APPLICANT:  as represented by Ther Secretary of the
;  APPLICANT:  Department of Health and Human Services
;  TITLE OF INVENTION: Mammalian T1R3 Sweet Taste Receptors
;  FILE REFERENCE: 02307E-140800US
;  CURRENT APPLICATION NUMBER: US/10/679,102
;  CURRENT FILING DATE:  2003-10-02
;  NUMBER OF SEQ ID NOS: 32
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
;   LENGTH: 852
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: human T1R3 G-protein coupled receptor sweet taste
;   OTHER INFORMATION: receptor
US-10-679-102-15

  Query Match           100.0%;  Score 4524;  DB 5;  Length 852;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  852;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360

Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420

Qy        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540

Qy        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600

Qy        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720

Qy        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780

Qy        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840

Qy        841 NDGNTGNQGKHE 852
                 ||||||||||||
Db        841 NDGNTGNQGKHE 852


RESULT 8
US-10-292-798-602
; Sequence 602, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
;  APPLICANT: SUWA, MAKIKO
;  APPLICANT:  ASAI, KIYOSHI
;  APPLICANT:  AKIYAMA, YUTAKA
;  APPLICANT:  ABURATANI, HIROYUKI
;  TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
;  FILE REFERENCE: 084335/166
;  CURRENT APPLICATION NUMBER: US/10/292,798
;  CURRENT FILING DATE:  2002-11-13
;  PRIOR APPLICATION NUMBER: 10/017,161
;  PRIOR FILING DATE: 2001-12-18
;  PRIOR APPLICATION NUMBER: JP 2001-246789
;  PRIOR FILING DATE: 2001-06-18
;  NUMBER OF SEQ ID NOS: 2070
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 602
;   LENGTH: 936
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-292-798-602

  Query Match            100.0%;  Score 4524;  DB 4;  Length 936;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  852;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360

Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420

Qy        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540

Qy        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600

Qy        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720

Qy        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780

Qy        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840

Qy        841 NDGNTGNQGKHE 852
                 ||||||||||||
Db        841 NDGNTGNQGKHE 852


RESULT 9
US-10-188-186-148
; Sequence 148, Application US/10188186
; Publication No. US20040029789A1
; GENERAL INFORMATION:
;  APPLICANT: Anderson et al.
;  TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-397C
;  CURRENT APPLICATION NUMBER: US/10/188,186
;  CURRENT FILING DATE:  2002-07-02
;  PRIOR APPLICATION NUMBER: 60/303046
;  PRIOR FILING DATE: 2001-07-05
;  PRIOR APPLICATION NUMBER: 60/360814
;  PRIOR FILING DATE: 2002-03-01
;  PRIOR APPLICATION NUMBER: 60/303828
;  PRIOR FILING DATE: 2001-09-07
;  PRIOR APPLICATION NUMBER: 60/323380
;  PRIOR FILING DATE: 2001-09-19
;  PRIOR APPLICATION NUMBER: 60/361133
;  PRIOR FILING DATE: 2002-03-01
;  PRIOR APPLICATION NUMBER: 60/304016
;  PRIOR FILING DATE: 2001-07-09
;  PRIOR APPLICATION NUMBER: 60/304502
;  PRIOR FILING DATE: 2001-07-11
;  PRIOR APPLICATION NUMBER: 60/305262
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: 60/373881
;  PRIOR FILING DATE: 2002-04-19
;  PRIOR APPLICATION NUMBER: 60/305673
;  PRIOR FILING DATE: 2001-07-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 368
;  SOFTWARE: Custom
; SEQ ID NO 148
;   LENGTH: 852
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-188-186-148

  Query Match            99.8%;  Score 4517;  DB 4;  Length 852;
```

```
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  851;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360

Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420

Qy        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERLKIRWHTSDNQVPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540

Qy        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600

Qy        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720

Qy        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780

Qy        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840

Qy        841 NDGNTGNQGKHE 852
              ||||||||||||
Db        841 NDGNTGNQGKHE 852


RESULT 10
US-09-796-338A-14
; Sequence 14, Application US/09796338A
; Patent No. US20020061522A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals, Inc.
;  TITLE OF INVENTION: 1983, 52881, 2398, 45449, 50289, AND
;  TITLE OF INVENTION:  52872, NOVEL G PROTEIN-COUPLED RECEPTORS AND USES THEREFOR
;  FILE REFERENCE: 10448-020001
;  CURRENT APPLICATION NUMBER: US/09/796,338A
;  CURRENT FILING DATE:  2001-02-28
;  PRIOR APPLICATION NUMBER: US 60/186,059
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
;   LENGTH: 852
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-796-338A-14

  Query Match             99.7%;  Score 4512;  DB 3;  Length 852;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  851;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360

Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420

Qy        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540

Qy        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600

Qy        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720

Qy        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
              |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGRYNRARGLTFAMLAYFITWVSFVP 780

Qy        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840

Qy        841 NDGNTGNQGKHE 852
```

```
                    ||||||||||||
Db        841 NDGNTGNQGKHE 852


RESULT 11
US-10-261-482-2
; Sequence 2, Application US/10261482
; Publication No. US20030036089A1
; GENERAL INFORMATION:
;  APPLICANT: WEI, Ming-Hui et al
;  TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
;  TITLE OF INVENTION:  RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
;  TITLE OF INVENTION:  PROTEINS, AND USES THEREOF
;  FILE REFERENCE: CL000869CON
;  CURRENT APPLICATION NUMBER: US/10/261,482
;  CURRENT FILING DATE:  2002-10-02
;  PRIOR APPLICATION NUMBER: 09/684,393
;  PRIOR FILING DATE: 2000-10-10
;  PRIOR APPLICATION NUMBER: 60/172,600
;  PRIOR FILING DATE: 1999-12-20
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 852
;   TYPE: PRT
;   ORGANISM: HUMAN
US-10-261-482-2

  Query Match             99.7%;  Score 4512;  DB 4;  Length 852;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  851;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360

Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420

Qy        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540

Qy        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600

Qy        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720

Qy        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGRYNRARGLTFAMLAYFITWVSFVP 780

Qy        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840

Qy        841 NDGNTGNQGKHE 852
              ||||||||||||
Db        841 NDGNTGNQGKHE 852


RESULT 12
US-10-282-837-14
; Sequence 14, Application US/10282837
; Publication No. US20030082738A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals, Inc.
;  TITLE OF INVENTION: 1983, 52881, 2398, 45449, 50289, AND
;  TITLE OF INVENTION:  52872, NOVEL G PROTEIN-COUPLED RECEPTORS AND USES THEREFOR
;  FILE REFERENCE: 10448-020001
;  CURRENT APPLICATION NUMBER: US/10/282,837
;  CURRENT FILING DATE:  2002-10-29
;  PRIOR APPLICATION NUMBER: US/09/796,338
;  PRIOR FILING DATE: 2001-02-28
;  PRIOR APPLICATION NUMBER: US 60/186,059
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
;   LENGTH: 852
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-282-837-14

  Query Match             99.7%;  Score 4512;  DB 4;  Length 852;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  851;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360

Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420

Qy        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540

Qy        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600

Qy        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720

Qy        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGRYNRARGLTFAMLAYFITWVSFVP 780

Qy        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840

Qy        841 NDGNTGNQGKHE 852
              ||||||||||||
Db        841 NDGNTGNQGKHE 852


RESULT 13
US-10-145-586-14
; Sequence 14, Application US/10145586
; Publication No. US20030138890A1
; GENERAL INFORMATION:
;  APPLICANT: Alexandra Glucksmann, Maria
;  APPLICANT:  Silos-Santiago, Inmaculada
;  APPLICANT:  M. Galvin, Katherine
;  APPLICANT:  Weich, Nadine
;  APPLICANT:  Curtis, Rory A.J.
;  APPLICANT:  Bandaru, Rajasekhar
;  APPLICANT:  Kapeller-Libermann, Rosana
;  TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTOR FAMILY MEMBERS,
;  TITLE OF INVENTION:  HUMAN THIOREDOXIN FAMILY MEMBERS, HUMAN LEUCINE-RICH
;  TITLE OF INVENTION:  REPEAT FAMILY MEMBERS, AND HUMAN RINGFINGER FAMILY MEMBER
;  FILE REFERENCE: 10448-188001
;  CURRENT APPLICATION NUMBER: US/10/145,586
;  CURRENT FILING DATE:  2002-05-14
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 95
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
;   LENGTH: 852
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-145-586-14

  Query Match             99.7%;  Score 4512;  DB 4;  Length 852;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  851;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360

Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420

Qy        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540

Qy        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600

Qy        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720

Qy        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGRYNRARGLTFAMLAYFITWVSFVP 780

Qy        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840

Qy        841 NDGNTGNQGKHE 852
              ||||||||||||
Db        841 NDGNTGNQGKHE 852


RESULT 14
US-10-407-079-90
; Sequence 90, Application US/10407079
; Publication No. US20030215860A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals, Inc.
;  APPLICANT:  Glucksmann, Maria A.
;  APPLICANT:  Silos-Santiago, Inmaculada
;  APPLICANT:  Carroll, Joseph M.
;  APPLICANT:  Galvin, Katherine M.
;  TITLE OF INVENTION: 18636, 2466, 43238, 1983, 52881, 2398,
;  TITLE OF INVENTION:  45449, 50289, 52872 AND 26908 MOLECULES AND USES THEREFOR
;  FILE REFERENCE: MPI03-051OMNIM
;  CURRENT APPLICATION NUMBER: US/10/407,079
;  CURRENT FILING DATE:  2003-04-03
;  PRIOR APPLICATION NUMBER: US 10/226,102
;  PRIOR FILING DATE: 2002-08-22
;  PRIOR APPLICATION NUMBER: US 60/314,041
;  PRIOR FILING DATE: 2001-08-22
```

```
;  PRIOR APPLICATION NUMBER: US 10/225,094
;  PRIOR FILING DATE: 2002-08-21
;  PRIOR APPLICATION NUMBER: US 60/314,185
;  PRIOR FILING DATE: 2001-08-22
;  PRIOR APPLICATION NUMBER: US 10/272,417
;  PRIOR FILING DATE: 2002-10-15
;  PRIOR APPLICATION NUMBER: US 09/715,790
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: US 60/191,845
;  PRIOR FILING DATE: 2000-03-24
;  PRIOR APPLICATION NUMBER: US 10/282,837
;  PRIOR FILING DATE: 2002-10-29
;  PRIOR APPLICATION NUMBER: US 09/796,338
;  PRIOR FILING DATE: 2001-02-28
;  PRIOR APPLICATION NUMBER: US 60/186,059
;  PRIOR FILING DATE: 2000-02-29
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 90
;   LENGTH: 852
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-407-079-90

  Query Match             99.7%;  Score 4512;  DB 4;  Length 852;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  851;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360

Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420

Qy        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540

Qy        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600

Qy        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720

Qy        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGRYNRARGLTFAMLAYFITWVSFVP 780

Qy        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840

Qy        841 NDGNTGNQGKHE 852
              ||||||||||||
Db        841 NDGNTGNQGKHE 852


RESULT 15
US-10-179-373-7
; Sequence 7, Application US/10179373
; Publication No. US20030232407A1
; GENERAL INFORMATION:
;  APPLICANT: ZOLLER, MARK
;  APPLICANT:  LI, XIAODONG
;  APPLICANT:  STASZEWSKI, LENA
;  APPLICANT:  O'CONNELL, SHAWN
;  APPLICANT:  ZOZULYA, SERGEY
;  APPLICANT:  ADLER, JON
;  APPLICANT:  XU, HONG
;  APPLICANT:  ECHEVERRI, FERNANDO
;  TITLE OF INVENTION: T1R HETERO-OLIGOMERIC TASTE RECEPTORS AND CELL LINES
;  TITLE OF INVENTION:  THAT EXPRESS SAID RECEPTORS AND USE THEREOF FOR
;  TITLE OF INVENTION:  IDENTIFICATION OF TASTE COMPOUNDS
;  FILE REFERENCE: 078003-0291566
;  CURRENT APPLICATION NUMBER: US/10/179,373
;  CURRENT FILING DATE:  2002-06-26
;  PRIOR APPLICATION NUMBER: 60/300,434
;  PRIOR FILING DATE: 2001-06-26
;  PRIOR APPLICATION NUMBER: 60/304,749
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: 60/310,493
;  PRIOR FILING DATE: 2001-08-08
;  PRIOR APPLICATION NUMBER: 60/331,771
;  PRIOR FILING DATE: 2001-11-21
;  PRIOR APPLICATION NUMBER: 60/339,472
;  PRIOR FILING DATE: 2001-12-14
;  PRIOR APPLICATION NUMBER: 60/372,090
;  PRIOR FILING DATE: 2002-04-15
;  PRIOR APPLICATION NUMBER: 60/374,143
;  PRIOR FILING DATE: 2002-04-22
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
;   LENGTH: 852
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-179-373-7

  Query Match             99.7%;  Score 4512;  DB 4;  Length 852;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  851;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
```

```
Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360

Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420

Qy        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540

Qy        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600

Qy        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720

Qy        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGRYNRARGLTFAMLAYFITWVSFVP 780

Qy        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840

Qy        841 NDGNTGNQGKHE 852
              ||||||||||||
Db        841 NDGNTGNQGKHE 852
```

Search completed: December 16, 2005, 18:15:16
Job time : 174 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2005, 18:06:46 ; Search time 12 Seconds
(without alignments)
478.311 Million cell updates/sec

Title: US-10-679-102-15
Perfect score: 4524
Sequence: 1 MLGPAVLGLSLWALLHPGTG.....GPGDAQGQNDGNTGNQGKHE 852

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 51470 seqs, 6736768 residues

Total number of hits satisfying chosen parameters: 51470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_New:*
1: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 4524 | 100.0 | 852 | 6 | US-10-645-441-15 | Sequence 15, Appl |
| 2 | 3255 | 71.9 | 858 | 6 | US-10-645-441-25 | Sequence 25, Appl |
| 3 | 3229 | 71.4 | 858 | 6 | US-10-645-441-23 | Sequence 23, Appl |
| 4 | 3227.5 | 71.3 | 858 | 6 | US-10-645-441-18 | Sequence 18, Appl |
| 5 | 3221.5 | 71.2 | 858 | 6 | US-10-645-441-20 | Sequence 20, Appl |
| 6 | 1208.5 | 26.7 | 840 | 6 | US-10-645-441-1 | Sequence 1, Appli |
| 7 | 1188.5 | 26.3 | 842 | 6 | US-10-645-441-2 | Sequence 2, Appli |
| 8 | 1095.5 | 24.2 | 928 | 6 | US-10-841-129-4 | Sequence 4, Appli |
| 9 | 1086.5 | 24.0 | 928 | 6 | US-10-841-129-6 | Sequence 6, Appli |
| 10 | 1061.5 | 23.5 | 926 | 6 | US-10-841-129-2 | Sequence 2, Appli |
| 11 | 1053 | 23.3 | 777 | 6 | US-10-645-441-3 | Sequence 3, Appli |
| 12 | 1011 | 22.3 | 838 | 6 | US-10-645-441-9 | Sequence 9, Appli |
| 13 | 1002.5 | 22.2 | 843 | 6 | US-10-645-441-7 | Sequence 7, Appli |
| 14 | 999.5 | 22.1 | 843 | 6 | US-10-645-441-8 | Sequence 8, Appli |
| 15 | 96.5 | 2.1 | 2919 | 6 | US-10-821-234-1133 | Sequence 1133, Ap |
| 16 | 95.5 | 2.1 | 389 | 6 | US-10-980-388-116 | Sequence 116, App |
| 17 | 95 | 2.1 | 1061 | 7 | US-11-059-814-18 | Sequence 18, Appl |
| 18 | 94.5 | 2.1 | 833 | 7 | US-11-076-187-5 | Sequence 5, Appli |
| 19 | 94.5 | 2.1 | 3655 | 7 | US-11-075-185-5 | Sequence 5, Appli |
| 20 | 94 | 2.1 | 705 | 7 | US-11-102-240-162 | Sequence 162, App |
| 21 | 94 | 2.1 | 7968 | 7 | US-11-143-980-49 | Sequence 49, Appl |
| 22 | 93 | 2.1 | 358 | 6 | US-10-980-388-96 | Sequence 96, Appl |
| 23 | 92 | 2.0 | 672 | 6 | US-10-467-657-8280 | Sequence 8280, Ap |
| 24 | 90 | 2.0 | 693 | 6 | US-10-467-657-6176 | Sequence 6176, Ap |
| 25 | 90 | 2.0 | 945 | 6 | US-10-131-826A-146 | Sequence 146, App |
| 26 | 90 | 2.0 | 2764 | 6 | US-10-995-561-691 | Sequence 691, App |
| 27 | 90 | 2.0 | 2813 | 6 | US-10-995-561-688 | Sequence 688, App |
| 28 | 89 | 2.0 | 617 | 7 | US-11-143-980-35 | Sequence 35, Appl |
| 29 | 87 | 1.9 | 223 | 6 | US-10-467-657-3974 | Sequence 3974, Ap |
| 30 | 87 | 1.9 | 1236 | 7 | US-11-080-991-68 | Sequence 68, Appl |
| 31 | 86.5 | 1.9 | 2280 | 7 | US-11-022-562-211 | Sequence 211, App |
| 32 | 86 | 1.9 | 693 | 6 | US-10-131-826A-406 | Sequence 406, App |
| 33 | 85 | 1.9 | 163 | 6 | US-10-821-234-1356 | Sequence 1356, Ap |
| 34 | 85 | 1.9 | 163 | 6 | US-10-965-694-1 | Sequence 1, Appli |
| 35 | 85 | 1.9 | 832 | 6 | US-10-512-109-29 | Sequence 29, Appl |
| 36 | 84.5 | 1.9 | 1798 | 6 | US-10-995-561-1033 | Sequence 1033, Ap |
| 37 | 84.5 | 1.9 | 1798 | 6 | US-10-995-561-1034 | Sequence 1034, Ap |
| 38 | 84 | 1.9 | 1070 | 7 | US-11-000-463-721 | Sequence 721, App |
| 39 | 83.5 | 1.8 | 355 | 6 | US-10-967-527A-22 | Sequence 22, Appl |
| 40 | 83 | 1.8 | 489 | 6 | US-10-793-626-2632 | Sequence 2632, Ap |
| 41 | 83 | 1.8 | 829 | 6 | US-10-512-109-27 | Sequence 27, Appl |
| 42 | 82.5 | 1.8 | 655 | 7 | US-11-184-860-1 | Sequence 1, Appli |
| 43 | 82.5 | 1.8 | 985 | 7 | US-11-113-424-61 | Sequence 61, Appl |
| 44 | 81.5 | 1.8 | 334 | 6 | US-10-858-730-121 | Sequence 121, App |
| 45 | 81 | 1.8 | 161 | 6 | US-10-965-694-12 | Sequence 12, Appl |

ALIGNMENTS

RESULT 1
US-10-645-441-15
; Sequence 15, Application US/10645441
; Publication No. US20050260599A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Ryba, Nicholas J.P.
; APPLICANT: Nelson, Greg
; APPLICANT: Hoon, Mark A.
; APPLICANT: Chandrashekar, Jayaram
; APPLICANT: Zhang, Yifeng
; APPLICANT: The Regents of the University of California
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by the Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: Mammalian Sweet Taste Receptors
; FILE REFERENCE: 02307E-120110US
; CURRENT APPLICATION NUMBER: US/10/645,441
; CURRENT FILING DATE: 2003-08-20
; PRIOR APPLICATION NUMBER: US/09/927,315
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: US 60/302,898
; PRIOR FILING DATE: 2001-07-03
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 852
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human T1R3 sweet taste receptor
US-10-645-441-15

Query Match 100.0%; Score 4524; DB 6; Length 852;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 852; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
```

```
Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360

Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420

Qy        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540

Qy        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600

Qy        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720

Qy        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780

Qy        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840

Qy        841 NDGNTGNQGKHE 852
              ||||||||||||
Db        841 NDGNTGNQGKHE 852
```

RESULT 2
US-10-645-441-25
; Sequence 25, Application US/10645441
; Publication No. US20050260599A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Ryba, Nicholas J.P.
; APPLICANT: Nelson, Greg
; APPLICANT: Hoon, Mark A.
; APPLICANT: Chandrashekar, Jayaram
; APPLICANT: Zhang, Yifeng
; APPLICANT: The Regents of the University of California
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by the Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: Mammalian Sweet Taste Receptors
; FILE REFERENCE: 02307E-120110US
; CURRENT APPLICATION NUMBER: US/10/645,441
; CURRENT FILING DATE: 2003-08-20
; PRIOR APPLICATION NUMBER: US/09/927,315
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: US 60/302,898
; PRIOR FILING DATE: 2001-07-03
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 858
; TYPE: PRT
; ORGANISM: Rattus sp.
; FEATURE:
; OTHER INFORMATION: rat T1R3 sweet taste receptor
US-10-645-441-25

Query Match 71.9%; Score 3255; DB 6; Length 858;
Best Local Similarity 72.9%; Pred. No. 7.4e-253;
Matches 623; Conservative 74; Mismatches 144; Indels 14; Gaps 4;

```
Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              | | |:||||| | |  | |: |||||| : :|||:||||||||  ||| |  ||:|:
Db          1 MPGLAILGLSLAAFLELGMGSSLCLSQQFKAQGDYILGGLFPLGTTEEATLNQRTQPNGI 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
              :|||||  ||  |:|||||||||||| | ||||||||||||||||||||| |||||||:||
Db         61 LCTRFSPLGLFLAMAMKMAVEEINNGSALLPGLRLGYDLFDTCSEPVVTMKPSLMFMAKV 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              ||: |||||||||||||||||||||::|||||||||||||| |||: || ||||
Db        121 GSQSIAAYCNYTQYQPRVLAVIGPHSSELALITGKFFSFFLMPQVSYSASMDRLSDRETF 180

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              |||||||||||||  |   ||| | ||||||||||:|||:|||||| || :||||||||
Db        181 PSFFRTVPSDRVQLQAVVTLLQNFSWNWVAALGSDDDYGREGLSIFSGLANSRGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||        :|||| ||| |||||  ||||:||||  | ::||:|||   |||||||||
Db        241 GLVPQHDTSGQQLGKVVDVLRQVNQSKVQVVVLFASARAVYSLFSYSILHDLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              |:|||||||| || :|::|||||||||||| | ||  ||:| |||| || ||::| : |
Db        301 ESWLTSDLVMTLPNIARVGTVLGFLQRGALLPEFSHYVETRLALAADPTFCASL-KAELD 359

Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLN--------HHQTFSVYAAVYSVAQALHNTLQCNA 412
              ||| |:| || ||| | |||:|:||          ||| |: ||||||||||||||||
Db        360 LEERVMGPRCSQCDYIMLQNLSSGLMQNLSAGQLHHQIFATYAAVYSVAQALHNTLQCNV 419

Qy        413 SGCPAQDPVKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVG 472
              | |   :||:|||||||||::|    | |:||: |:||||||:||||   | || ||
Db        420 SHCHTSEPVQPWQLLENMYNMSFRARDLTLQFDAKGSVDMEYDLKMWVWQSPTPVLHTVG 479

Qy        473 RFNGSLRTERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQN 532
              |||:|: :  |: |    || |||:|||||::|||||||||||||||:|||||::
Db        480 TFNGTLQLQHSKMYW--PGNQVPVSQCSRQCKDGQVRRVKGFHSCCYDCVDCKAGSYRKH 537

Qy        533 PDDIACTFCGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVH 592
              |||  || ||:|:||||:|| |  || :||||||||| ||||| | ||| |||||||||
Db        538 PDDFTCTPCGKDQWSPEKSTTCLPRRPKFLAWGEPAVLSLLLLLCLVLGLTLAALGLFVH 597

Qy        593 HRDSPLVQASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLF 652
              : |||||||||| | ||||:|||| |||||||||:|  | |||||||::|||||||||||
Db        598 YWDSPLVQASGGSLFCFGLICLGLFCLSVLLFPGRPRSASCLAQQPMAHLPLTGCLSTLF 657

Qy        653 LQAAEIFVESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWH 712
              |||||||||||||||: |   |||||||||||| ||| ||| |||:|||||||||
Db        658 LQAAEIFVESELPLSWANWLCSYLRGPWAWLVVLLATLVEAALCAWYLMAFPPEVVTDWQ 717

Qy        713 MLPTEALVHCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYF 772
              :|||| | ||| ||||| || | ||| |||||||||||||:|||| ||||||||||||||
Db        718 VLPTEVLEHCRMRSWVSLGLVHITNAVLAFLCFLGTFLVQSQPGRYNRARGLTFAMLAYF 777

Qy        773 ITWVSFVPLLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGG 832
              | |||||||||||||  :||||||:| | ||||| |||||:||:|:  | ||| |||||
Db        778 IIWVSFVPLLANVQVAYQPAVQMGAILFCALGILATFHLPKCYVLLWLPELNTQEFFLGR 837

Qy        833 GPGDAQGQNDGNTGN 847
              | :|   :|||:|:
Db        838 SPKEA---SDGNSGS 849
```

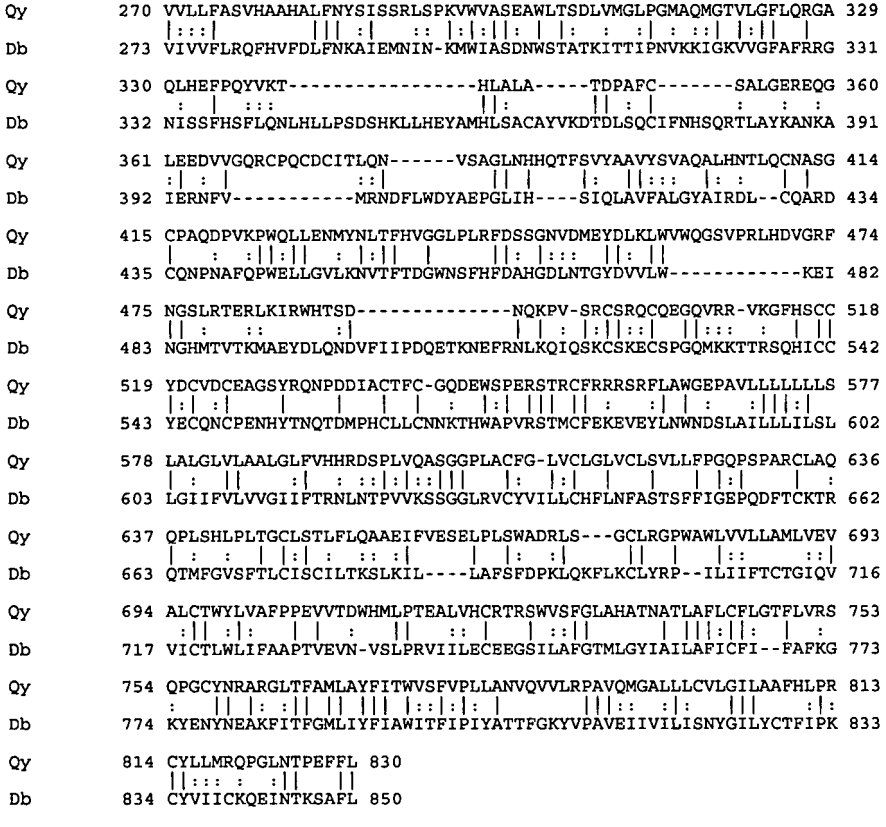
```
RESULT 3
US-10-645-441-23
; Sequence 23, Application US/10645441
; Publication No. US20050260599A1
; GENERAL INFORMATION:
;  APPLICANT: Zuker, Charles S.
;  APPLICANT:  Ryba, Nicholas J.P.
;  APPLICANT:  Nelson, Greg
;  APPLICANT:  Hoon, Mark A.
;  APPLICANT:  Chandrashekar, Jayaram
;  APPLICANT:  Zhang, Yifeng
;  APPLICANT:  The Regents of the University of California
;  APPLICANT:  The Government of the United States of America
;  APPLICANT:  as represented by the Secretary of the
;  APPLICANT:  Department of Health and Human Services
;  TITLE OF INVENTION: Mammalian Sweet Taste Receptors
;  FILE REFERENCE: 02307E-120110US
;  CURRENT APPLICATION NUMBER: US/10/645,441
;  CURRENT FILING DATE:  2003-08-20
;  PRIOR APPLICATION NUMBER: US/09/927,315
;  PRIOR FILING DATE: 2001-08-10
;  PRIOR APPLICATION NUMBER: US 60/302,898
;  PRIOR FILING DATE: 2001-07-03
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
;   LENGTH: 858
;   TYPE: PRT
;   ORGANISM: Mus musculus
;   FEATURE:
;   OTHER INFORMATION: mouse Sac taster C57 T1R3 sweet taste receptor
US-10-645-441-23

  Query Match             71.4%;  Score 3229;  DB 6;  Length 858;
  Best Local Similarity   72.5%;  Pred. No. 8.9e-251;
  Matches  621;  Conservative   67;  Mismatches  147;  Indels   22;  Gaps    4;

Qy          5 AVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSPVCTR 64
              |::|||| | |   | || |||||| : :|||:||||||||  ||| |  ||:|:|  | |
Db          5 AIMGLSLAAFLELGMGASLCLSQQPKAQGDYILGGLFPLGSTEEATLNQRTQPNSIPCNR 64

Qy         65 FSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKAGSRD 124
              ||  ||  |:||||||||||| | ||||||||||||||||||| || ||||||| ||:
Db         65 FSPLGLFLAMAMKMAVEEINNGSALLPGLRLGYDLFDTCSEPVVTMKSSLMFLAKVGSQS 124

Qy        125 IAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETFPSFF 184
              |||||||||||||||||||||||::|||||||||||||| |||: || ||||||||
Db        125 IAAYCNYTQYQPRVLAVIGPHSSELALITGKFFSFFLMPQVSYSASMDRLSDRETFPSFF 184

Qy        185 RTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHEGLVP 244
              |||||||||| |   ||| | |||||||||||:|||:||||||:|| ||||||||||||
Db        185 RTVPSDRVQLQAVVTLLQNFSWNWVAALGSDDDYGREGLSIFSSLANARGICIAHEGLVP 244

Qy        245 LPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVASEAWL 304
                  :|||| ||| |||| ||||:||||  | ::||:|||  ||||||||||:||
Db        245 QHDTSGQQLGKVLDVLRQVNQSKVQVVVLFASARAVYSLFSYSIHHGLSPKVWVASESWL 304

Qy        305 TSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQGLEED 364
              |||||| || :|::|||||||||| | ||  ||:|||||| |||||::|   |  |||
Db        305 TSDLVMTLPNIARVGTVLGFLQRGALLPEFSHYVETHLALAADPAFCASL-NAELDLEEH 363

Qy        365 VVGQRCPQCDCITLQNVSAGL--------NHHQTFSVYAAVYSVAQALHNTLQCNASGCP 416
              |:|||||:|| | |||:|:||         ||| |: ||||||||||||||||| | |
Db        364 VMGQRCPRCDDIMLQNLSSGLLQNLSAGQLHHQIFATYAAVYSVAQALHNTLQCNVSHCH 423

Qy        417 AQDPVKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNG 476
              : | |||||||||::||   | |:||: |||||||||:||||   | || || |||
Db        424 VSEHVLPWQLLENMYNMSFHARDLTLQFDAEGNVDMEYDLKMWVWQSPTPVLHTVGTFNG 483

Qy        477 SLRTERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDI 536
              :|: :: |: |    || |||:|||||::|||||||||||||||||:|||||::|||
Db        484 TLQLQQSKMYW--PGNQVPVSQCSRQCKDGQVRRVKGFHSCCYDCVDCKAGSYRKHPDDF 541

Qy        537 ACTFCGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDS 596
              || | ||:||||:|| |   || :||||||| || ||||| | ||| ||||||| ||| ||
Db        542 TCTPCNQDQWSPEKSTACLPRRPKFLAWGEPVVLSLLLLLCLVLGLALAALGLSVHHWDS 601

Qy        597 PLVQASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAA 656
              ||||||||   ||||:|||| |||||||||:|| | |||||||::||||||||||||||
Db        602 PLVQASGGSQFCFGLICLGLFCLSVLLFPGRPSSASCLAQQPMAHLPLTGCLSTLFLQAA 661

Qy        657 EIFVESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPT 716
              | |||||||||||: |   ||| |||||||||  || ||| |||:|||||||||| :|||
Db        662 ETFVESELPLSWANWLCSYLRGLWAWLVVLLATFVEAALCAWYLIAFPPEVVTDWSVLPT 721

Qy        717 EALVHCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWV 776
              | | ||  ||||| || | ||| ||||||||||||:|||| ||||||||||||||||||
Db        722 EVLEHCHVRSWVSLGLVHITNAMLAFLCFLGTFLVQSQPGRYNRARGLTFAMLAYFITWV 781

Qy        777 SFVPLLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLG----- 831
              |||||||||||  :|||||||:|:| ||||  ||||:||:|:  | ||| |||||
Db        782 SFVPLLANVQVAYQPAVQMGAILVCALGILVTFHLPKCYVLLWLPKLNTQEFFLGRNAKK 841

Qy        832 ------GGPGDAQGQND 842
                     ||   ||| |:
Db        842 AADENSGGGEAAQGHNE 858


RESULT 4
US-10-645-441-18
; Sequence 18, Application US/10645441
; Publication No. US20050260599A1
; GENERAL INFORMATION:
;  APPLICANT: Zuker, Charles S.
;  APPLICANT:  Ryba, Nicholas J.P.
;  APPLICANT:  Nelson, Greg
;  APPLICANT:  Hoon, Mark A.
;  APPLICANT:  Chandrashekar, Jayaram
;  APPLICANT:  Zhang, Yifeng
;  APPLICANT:  The Regents of the University of California
;  APPLICANT:  The Government of the United States of America
;  APPLICANT:  as represented by the Secretary of the
;  APPLICANT:  Department of Health and Human Services
;  TITLE OF INVENTION: Mammalian Sweet Taste Receptors
;  FILE REFERENCE: 02307E-120110US
;  CURRENT APPLICATION NUMBER: US/10/645,441
;  CURRENT FILING DATE:  2003-08-20
;  PRIOR APPLICATION NUMBER: US/09/927,315
;  PRIOR FILING DATE: 2001-08-10
;  PRIOR APPLICATION NUMBER: US 60/302,898
;  PRIOR FILING DATE: 2001-07-03
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
;   LENGTH: 858
;   TYPE: PRT
;   ORGANISM: Mus musculus
;   FEATURE:
;   OTHER INFORMATION: mouse Sac non taster 129 T1R3 sweet taste receptor
US-10-645-441-18

  Query Match             71.3%;  Score 3227.5;  DB 6;  Length 858;
  Best Local Similarity   72.4%;  Pred. No. 1.2e-250;
  Matches  620;  Conservative   67;  Mismatches  156;  Indels   13;  Gaps    4;

Qy          5 AVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSPVCTR 64
              |::|||| | |   | || |||||| : :|||:||||||||  ||| |  | :|:| :| |
Db          5 AIMGLSLAAFLELGMGASLCLSQQFKAQGDYILGGLFPLGSTEEATLNQRAQPNSTLCNR 64

Qy         65 FSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKAGSRD 124
              ||  ||  |:||||||||||| | ||||||||||||||||||| || ||||||| ||:
Db         65 FSPLGLFLAMAMKMAVEEINNGSALLPGLRLGYDLFDTCSEPVVTMKSSLMFLAKVGSQS 124

Qy        125 IAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETFPSFF 184
```

```
Db        125 IAAYCNYTQYQPRVLAVIGPHSSELALITGKFFSFFLMPQVSYSASMDRLSDRETFPSFF 184

Qy        185 RTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHEGLVP 244

Db        185 RTVPSDRVQLQAVVTLLQNFSWNWVAALGSDDDYGREGLSIFSSLANARGICIAHEGLVP 244

Qy        245 LPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVASEAWL 304

Db        245 QHDTSGQQLGKVLDVLRQVNQSKVQVVVLFASARAVYSLFSYSIHHGLSPKVWVASESWL 304

Qy        305 TSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQGLEED 364

Db        305 TSDLVMTLPNIARVGTVLGFLQRGALLPEFSHYVETHLALAADPAFCASL-NAELDLEEH 363

Qy        365 VVGQRCPQCDCITLQNVSAGL--------NHHQTFSVYAAVYSVAQALHNTLQCNASGCP 416

Db        364 VMGQRCPQCDDIMLQNLSSGLLQNLSAGQLHHQIFATYAAVYSVAQALHNTLQCNVSHCH 423

Qy        417 AQDPVKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNG 476

Db        424 VSEHVLPWQLLENMYNMSFHARDLTLQFDAEGNVDMEYDLKMWVWQSPTPVLHTVGTFNG 483

Qy        477 SLRTERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDI 536

Db        484 TLQLQQSKMYW--PGNQVPVSQCSRQCKDGQVRRVKGFHSCCYDCVDCKAGSYRKHPDDF 541

Qy        537 ACTFCGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDS 596

Db        542 TCTPCNQDQWSPEKSTACLPRRPKFLAWGEPVVLSLLLLLCLVLGLALAALGLSVHHWDS 601

Qy        597 PLVQASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAA 656

Db        602 PLVQASGGSQFCFGLICLGLFCLSVLLFPGRPSSASCLAQQPMAHLPLTGCLSTLFLQAA 661

Qy        657 EIFVESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPT 716

Db        662 ETFVESELPLSWANWLCSYLRGLWAWLVVLLATFVEAALCAWYLTAFPPEVVTDWSVLPT 721

Qy        717 EALVHCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWV 776

Db        722 EVLEHCHVRSWVSLGLVHITNAMLAFLCFLGTFLVQSQPGRYNRARGLTFAMLAYFITWV 781

Qy        777 SFVPLLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGD 836

Db        782 SFVPLLANVQVAYQPAVQMGAILVCALGILVTFHLPKCYVLLWLPKLNTQEFFLGRNAKK 841

Qy        837 AQGQNDGNTGNQGKHE 852

Db        842 AADENSG--GGEAAQE 855


RESULT 5
US-10-645-441-20
; Sequence 20, Application US/10645441
; Publication No. US20050260599A1
; GENERAL INFORMATION:
;  APPLICANT: Zuker, Charles S.
;  APPLICANT:  Ryba, Nicholas J.P.
;  APPLICANT:  Nelson, Greg
;  APPLICANT:  Hoon, Mark A.
;  APPLICANT:  Chandrashekar, Jayaram
;  APPLICANT:  Zhang, Yifeng
;  APPLICANT:  The Regents of the University of California
;  APPLICANT:  The Government of the United States of America
;  APPLICANT:  as represented by the Secretary of the
;  APPLICANT:  Department of Health and Human Services
;  TITLE OF INVENTION: Mammalian Sweet Taste Receptors
;  FILE REFERENCE: 02307E-120110US
;  CURRENT APPLICATION NUMBER: US/10/645,441
;  CURRENT FILING DATE:  2003-08-20
;  PRIOR APPLICATION NUMBER: US/09/927,315
;  PRIOR FILING DATE: 2001-08-10
;  PRIOR APPLICATION NUMBER: US 60/302,898
;  PRIOR FILING DATE: 2001-07-03
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
;   LENGTH: 858
;   TYPE: PRT
;   ORGANISM: Mus musculus
;   FEATURE:
;   OTHER INFORMATION: mouse Sac taster SWR T1R3 sweet taste receptor
US-10-645-441-20

  Query Match             71.2%;  Score 3221.5;  DB 6;  Length 858;
  Best Local Similarity   72.4%;  Pred. No. 3.5e-250;
  Matches  620;  Conservative   67;  Mismatches  156;  Indels   13;  Gaps    4;

Qy          5 AVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSPVCTR 64

Db          5 AIMGLSLAAFLELGMGASLCLSQQFKAQGDYILGGLFPLGSTEEATLNQRTQPNSILCNR 64

Qy         65 FSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKAGSRD 124

Db         65 FSPLGLFLAMAMKMAVEEINNGSALLPGLRLGYDLFDTCSEPVVTMKSSLMFLAKVGSQS 124

Qy        125 IAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETFPSFF 184

Db        125 IAAYCNYTQYQPRVLAVIGPHSSELALITGKFFSFFLMPQVSYSASMDRLSDRETFPSFF 184

Qy        185 RTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHEGLVP 244

Db        185 RTVPSDRVQLQAVVTLLQNFSWNWVAALGSDDDYGREGLSIFSSLANARGICIAHEGLVP 244

Qy        245 LPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVASEAWL 304

Db        245 QHDTSGQQLGKVLDVLCQVNQSKVQVVVLFASARAVYSLFSYSIHHGLSPKVWVASESWL 304

Qy        305 TSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQGLEED 364

Db        305 TSDLVMTLPNIARVGTVLGFLQRGALLPEFSHYVETHLALAADPAFCASL-NAELDLEEH 363

Qy        365 VVGQRCPQCDCITLQNVSAGL--------NHHQTFSVYAAVYSVAQALHNTLQCNASGCP 416

Db        364 VMGQRCPQCDDIMLQNLSSGLLQNLSAGQLHHQIFATYAAVYSVAQALHNTLQCNVSHCH 423

Qy        417 AQDPVKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNG 476

Db        424 VSEHVLPWQLLENMYNMSFHARDLTLQFDAEGNVDMEYDLKMWVWQSPTPVLHTVGTFNG 483

Qy        477 SLRTERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDI 536

Db        484 TLQLQQSKMYW--PGNQVPVSQCSRQCKDGQVRRVKGFHSCCYDCVDCKAGSYRKHPDDF 541

Qy        537 ACTFCGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDS 596

Db        542 TCTPCNQDQWSPEKSTACLPRRPKFLAWGEPVVLSLLLLLCLVLGLALAALGLSVHHWDS 601

Qy        597 PLVQASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAA 656

Db        602 PLVQASGGSQFCFGLICLGLFCLSVLLFPGRPSSASCLAQQPMAHLPLTGCLSTLFLQAA 661

Qy        657 EIFVESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPT 716

Db        662 ETFVESELPLSWANWLCSYLRGLWAWLVVLSATFVEAALCAWYLTAFPPEVVTDWSVLPT 721

Qy        717 EALVHCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWV 776

Db        722 EVLEHCHVRSWVSLGLVHITNAMLAFLCFLGTFLVQSQPGRYNRARGLTFAMLAYFITWV 781

Qy        777 SFVPLLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGD 836

Db        782 SFVPLLANVQVAYQPAVQMGAILVCALGILVTFHLPKCYVLLWLPKLNTQEFFLGRNAKK 841

Qy        837 AQGQNDGNTGNQGKHE 852
```

```
                |  :| |  | :   |
Db        842 AADENSG--GGEAAQE 855


RESULT 6
US-10-645-441-1
; Sequence 1, Application US/10645441
; Publication No. US20050260599A1
; GENERAL INFORMATION:
;  APPLICANT: Zuker, Charles S.
;  APPLICANT:  Ryba, Nicholas J.P.
;  APPLICANT:  Nelson, Greg
;  APPLICANT:  Hoon, Mark A.
;  APPLICANT:  Chandrashekar, Jayaram
;  APPLICANT:  Zhang, Yifeng
;  APPLICANT:  The Regents of the University of California
;  APPLICANT:  The Government of the United States of America
;  APPLICANT:  as represented by the Secretary of the
;  APPLICANT:  Department of Health and Human Services
;  TITLE OF INVENTION: Mammalian Sweet Taste Receptors
;  FILE REFERENCE: 02307E-120110US
;  CURRENT APPLICATION NUMBER: US/10/645,441
;  CURRENT FILING DATE:  2003-08-20
;  PRIOR APPLICATION NUMBER: US/09/927,315
;  PRIOR FILING DATE: 2001-08-10
;  PRIOR APPLICATION NUMBER: US 60/302,898
;  PRIOR FILING DATE: 2001-07-03
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 840
;   TYPE: PRT
;   ORGANISM: Rattus sp.
;   FEATURE:
;   OTHER INFORMATION: rat T1R1 sweet taste receptor
US-10-645-441-1

  Query Match             26.7%;  Score 1208.5;  DB 6;  Length 840;
  Best Local Similarity   33.8%;  Pred. No. 4e-89;
  Matches  291;  Conservative  151;  Mismatches  339;  Indels   79;  Gaps   18;

Qy         12 WALLHPGTGAPLCLSQQL-----------------RMKGDYVLGGLFPL-GEAEEAGLRS 53
              ||        | | || ||                 : ||::| ||| | |:     |:
Db          4 WA-------AHLLLSLQLVYCWAFSCQRTESSPGFSLPGDFLLAGLFSLHGDC----LQV 52

Qy         54 RTRPSSPVCTR---FSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAM 110
              | ||    | |   |: :|     ||:  |||||| | ||| : |||:|:| |||    :
Db         53 RHRPLVTSCDRPDSFNGHGYHLFQAMRFTVEEINNSSALLPNITLGYELYDVCSES-ANV 111

Qy        111 KPSLMFLAKAGSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGAS 170
              :|  ||  | | |    :  :  :|:| ||| ::: |: |       |||| ||| ||
Db        112 YATLRVLALQGPRHIEIQKDLRNHSSKVVAFIGPDNTDHAVTTAALLGPFLMPLVSYEAS 171

Qy        171 MELLSARETFPSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALA 230
              :|||:  |||| ||||||| |:    :||| ||| |:: :||  :||: |:     ||
Db        172 SVVLSAKRKFPSFLRTVPSDRHQVEVMVQLLQSFGWVWISLIGSYGDYGQLGVQALEELA 231

Qy        231 AARGICIAHEGLVPLPRADDSRLG--KVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSI 288
              ||||:| : :||      :|:|  ::| ::  : |:   ||::|:: | |   |   :
Db        232 VPRGICVAFKDIVPF----SARVGDPRMQSMMQHLAQARTTVVVVFSNRHLARVFFRSVV 287

Qy        289 SSRLSPKVWVASEAWLTSDLVMGLPGMAQMGTVLGFLQRGAQ---LHEFPQ-YVKTHLAL 344
              : |: |||||| |  |  :  : |:  :|||||   :  |   | || : ||:   |
Db        288 LANLTGKVWVASEDWAISTYITSVTGIQGIGTVLGVAVQQRQVPGLKEFEESYVRAVTAA 347

Qy        345 AT---DPAFCSALGEREQGLEEDVVGQRCPQCDCITLQNVS--AGLNHHQTFSVYAAVYS 399
              :   : ::||               | | :|   | :|:      :    : || |||:
Db        348 PSACPEGSWCS------------TNQLCRECHTFTTRNMPTLGAFSMSAAYRVYEAVYA 394

Qy        400 VAQALHNTLQCNASGCPAQDPVKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLW 459
              ||  ||  | | :  | :: || ||||||: :| : | :    : || :|:    ||:  |
Db        395 VAHGLHQLLGCTSEIC-SRGPVYPWQLLQQIYKVNFLLHENTVAFDDNGDTLGYYDIIAW 453

Qy        460 VWQGSVPRLHDVGRFNGS---LRTERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHS 516
              | |       :|  : |   |   : ||:||  :|| ||| |:  |  |  | | | |
Db        454 DWNGPEWTFEIIGSASLSPVHLDINKTKIQWHGKNNQVPVSVCTTDCLAGHHRVVVGSHH 513

Qy        517 CCYDCVDCEAGSYRQNPDDIACTFCGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLL 576
              ||::|| ||||::    :   |  || :||:|: || || |   |||| ||  |:|:
Db        514 CCFECVPCEAGTFLNMSELHICQPCGTEEWAPKESTTCFPRTVEFLAWHEPISLVLIAAN 573

Qy        577 SLALGLVLAALGLFVHHRDSPLVQASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQ 636
              :| | |::   |||  |  :|:|:::|| |    |  |      |  | |:|:   || :
Db        574 TLLLLLLVGTAGLFAWHFHTPVVRSAGGRLCFLMLGSLVAGSCSFYSFFGEPTVPACLLR 633

Qy        637 QPLSHLPLTGCLSTLFLQAAEIFV----ESELPL---SWADRLSGCLRGPWAWLVVLLAM 689
              |||  |      || | ::: :: :     :::|    :||       :   | | |:::
Db        634 QPLFSLGFAIFLSCLTIRSFQLVIIFKFSTKVPTFYRTWA-------QNHGAGLFVIVSS 686

Qy        690 LVEVALCTWYLVAFPPEVVTDWHMLPTEALVHCRTRSWVSFGLAHATNATLAFLCFLGTF 749
              | : :|  :|| : |    ::   |   :: |   : | | ||   |  |:   |: ::
Db        687 TVHLLICLTWLVMWTPRPTREYQRFPHLVILECTEVNSVGFLLAFTHNILLSISTFVCSY 746

Qy        750 LVRSQPGCYNRARGLTFAMLAYFITWVSFVPLLANVQVVLRPAVQMGALLLCVLGILAAF 809
              | :  |  || |: :||::|  |::|::|  : :  |   ||| : | |  : |  : : :
Db        747 LGKELPENYNEAKCVTFSLLLNFVSWIAFFTMASIYQGSYLPAVNVLAGLTTLSGGFSGY 806

Qy        810 HLPRCYLLMRQPGLNTPEFF 829
              ||:||::: :| ||  | |
Db        807 FLPKCYVILCRPELNNTEHF 826


RESULT 7
US-10-645-441-2
; Sequence 2, Application US/10645441
; Publication No. US20050260599A1
; GENERAL INFORMATION:
;  APPLICANT: Zuker, Charles S.
;  APPLICANT:  Ryba, Nicholas J.P.
;  APPLICANT:  Nelson, Greg
;  APPLICANT:  Hoon, Mark A.
;  APPLICANT:  Chandrashekar, Jayaram
;  APPLICANT:  Zhang, Yifeng
;  APPLICANT:  The Regents of the University of California
;  APPLICANT:  The Government of the United States of America
;  APPLICANT:  as represented by the Secretary of the
;  APPLICANT:  Department of Health and Human Services
;  TITLE OF INVENTION: Mammalian Sweet Taste Receptors
;  FILE REFERENCE: 02307E-120110US
;  CURRENT APPLICATION NUMBER: US/10/645,441
;  CURRENT FILING DATE:  2003-08-20
;  PRIOR APPLICATION NUMBER: US/09/927,315
;  PRIOR FILING DATE: 2001-08-10
;  PRIOR APPLICATION NUMBER: US 60/302,898
;  PRIOR FILING DATE: 2001-07-03
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 842
;   TYPE: PRT
;   ORGANISM: Mus musculus
;   FEATURE:
;   OTHER INFORMATION: mouse T1R1 sweet taste receptor
US-10-645-441-2

  Query Match             26.3%;  Score 1188.5;  DB 6;  Length 842;
  Best Local Similarity   33.1%;  Pred. No. 1.6e-87;
  Matches  284;  Conservative  159;  Mismatches  341;  Indels   75;  Gaps   17;

Qy         12 WALLHPGTGAPLCLSQQL-----------------RMKGDYVLGGLFPLGEAEEAGLR 52
              ||        | | || ||                 : ||::| ||| |        |:
Db          4 WA-------AHLLLSLQLAVAYCWAFSCQRTESSPGFSLPGDFLLAGLFSL---HADCLQ 53

Qy         53 SRTRPSSPVCTR---FSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEP--V 107
              | ||    | |   |: :|     ||:  |||||| : ||| : |||:|:| |||    |
Db         54 VRHRPLVTSCDRSDSFNGHGYHLFQAMRFTVEEINNSTALLPNITLGYELYDVCSESSNV 113
```

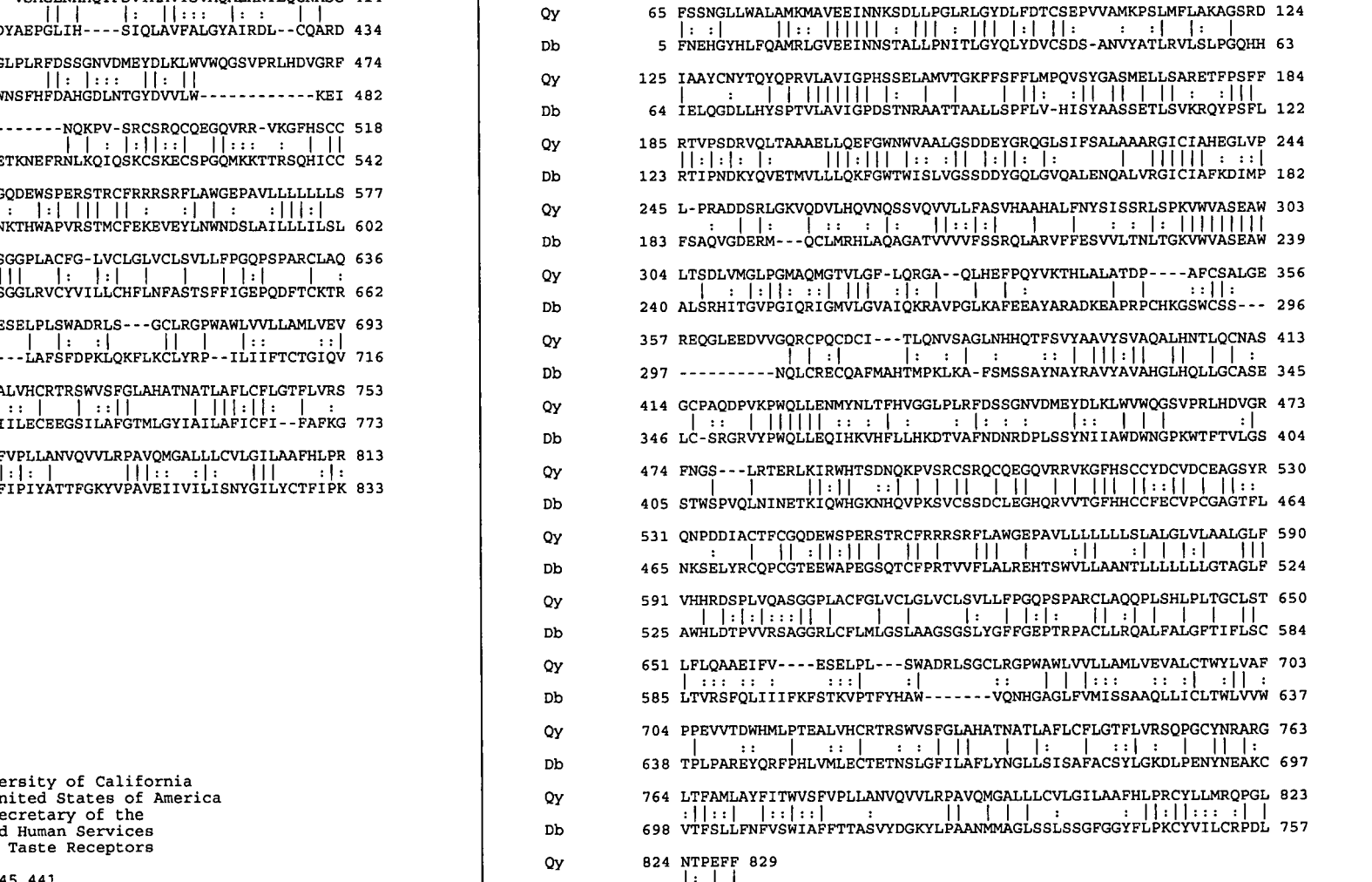
```
Qy        108 VAMKPSLMFLAKAGSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSY 167
Db        114 YA---TLRVPAQQGTGHLEMQRDLRNHSSKVVALIGPDNTDHAVTTAALLSPFLMPLVSY 170

Qy        168 GASMELLSARETFPSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFS 227
Db        171 EASSVILSGKRKFPSFLRTIPSDKYQVEVIVRLLQSFGWVWISLVGSYGDYGQLGVQALE 230

Qy        228 ALAAARGICIAHEGLVPL-PRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNY 286
Db        231 ELATPRGICVAFKDVVPLSAQAGDPRM---QRMMLRLARARTTVVVVFSNRHLAGVFFRS 287

Qy        287 SISSRLSPKVWVASEAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHL-ALA 345
Db        288 VVLANLTGKVWIASEDWAISTYITNVPGIQGIGTVLGVAIQQRQVPGLKEFEESYVQAVM 347

Qy        346 TDPAFC---SALGEREQGLEEDVVGQRCPQCDCITLQNVS--AGLNHHQTFSVYAAVYSV 400
Db        348 GAPRTCPEGSWCG----------TNQLCRECHAFTTWNMPELGAFSMSAAYNVYEAVYAV 397

Qy        401 AQALHNTLQCNASGCPAQDPVKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWV 460
Db        398 AHGLHQLLGCTSGTC-ARGPVYPWQLLQQIYKVNFLLHKKTVAFDDKGDPLGYYDIIAWD 456

Qy        461 WQGSVPRLHDVGRFNGS---LRTERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSC 517
Db        457 WNGPEWTFEVIGSASLSPVHLDINKTKIQWHGKNNQVPVSVCTRDCLEGHHRLVMGSHHC 516

Qy        518 CYDCVDCEAGSYRQNPDDIACTFCGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLS 577
Db        517 CFECMPCEAGTFLNTSELHTCQPCGTEEWAPEGSSACFSRTVEFLGWHEPISLVLLAANT 576

Qy        578 LALGLVLAALGLFVHHRDSPLVQASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQ 637
Db        577 LLLLLLIGTAGLFAWRLHTPVVRSAGGRLCFLMLGSLVAGSCSLYSFFGKPTVPACLLRQ 636

Qy        638 PLSHLPLTGCLSTLFLQAAEIFV----ESELPL---SWADRLSGCLRGPWAWLVVLLAML 690
Db        637 PLFSLGFAIFLSCLTIRSFQLVIIFKFSTKVPTFYHTWA-------QNHGAGIFVIVSST 689

Qy        691 VEVALCTWYLVAFPPEVVTDWHMLPTEALVHCRTRSWVSFGLAHATNATLAFLCFLGTFL 750
Db        690 VHLFLCLTWLAMWTPRPTREYQRFPHLVILECTEVNSVGFLVAFAHNILLSISTFVCSYL 749

Qy        751 VRSQPGCYNRARGLTFAMLAYFITWVSFVPLLANVQVVLRPAVQMGALLLCVLGILAAFH 810
Db        750 GKELPENYNEAKCVTFSLLLHFVSWIAFFTMSSIYQGSYLPAVNVLAGLATLSGGFSGYF 809

Qy        811 LPRCYLLMRQPGLNTPEFF 829
Db        810 LPKCYVILCRPELNNTEHF 828


RESULT 8
US-10-841-129-4
; Sequence 4, Application US/10841129
; Publication No. US20050250113A1
; GENERAL INFORMATION:
;  APPLICANT: Zuker, Charles S.
;  APPLICANT:  Erlenbach, Isolda
;  APPLICANT:  Hoon, Mark A.
;  APPLICANT:  Ryba, Nicholas J. P.
;  APPLICANT:  Zhang, Yifeng
;  APPLICANT:  The Regents of the University of California
;  APPLICANT:  The Government of the United States of America
;  APPLICANT:  as represented by The Secretary of the
;  APPLICANT:  Department of Health and Human Services
;  TITLE OF INVENTION: A Mammalian Magnesium/Manganese Sensing G Protein
;  TITLE OF INVENTION:  Coupled Receptor
;  FILE REFERENCE: 02307E-145400US
;  CURRENT APPLICATION NUMBER: US/10/841,129
;  CURRENT FILING DATE:  2004-05-07
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;  LENGTH: 928
;  TYPE: PRT
;  ORGANISM: Mus sp.
;  FEATURE:
;  OTHER INFORMATION: mouse magnesium/manganese sensing G protein
;  OTHER INFORMATION: coupled receptor (GPCR) R5.24, Mg receptor, Mn
;  OTHER INFORMATION: receptor
US-10-841-129-4

  Query Match             24.2%;  Score 1095.5;  DB 6;  Length 928;
  Best Local Similarity   31.0%;  Pred. No. 5e-80;
  Matches  263;  Conservative  161;  Mismatches  345;  Indels   79;  Gaps   21;

Qy         33 GDYVLGGLFPLGEAEEAGLRSRTRPSSPVCTRFSSNGLLWALAMKMAVEEINNKSDLLPG 92
Db         34 GHIMIGGLFAIHEKMLSSDDHPRRPQIQKCAGFEISVFLQTLAMIHSIEMINN-STLLSG 92

Qy         93 LRLGYDLFDTCSEPVVAMKPSLMFLAKAG-SRDIAAY-CNYTQYQPRVLAVIGPHSSELA 150
Db         93 VKLGYEIYDTCTEVTAAMAATLRFLSKFNCSRETVVFQCDYSSYMPRVKAVIGAGYSETS 152

Qy        151 MVTGKFFSFFLMPQVSYGASMELLSARETFPSFFRTVPSDRVQLTAAAELLQEFGWNWVA 210
Db        153 IAVSRMLNLQLMPQVSYESTAEILSDKIRFPSFLRTVPSDFYQTKAMAHLIRQSGWNWIG 212

Qy        211 ALGSDDEYGRQGLSIFSALAAARGICIAHEGLVPLPRADDSRLGKVQDVLHQ-VNQSSVQ 269
Db        213 AITTDDDYGRLALNTFAIQAAENNVCIAFKEVLPAFLSDNTIEVRINQTLEKIIAEAQVN 272

Qy        270 VVLLFASVHAAHALFNYSISSRLSPKVWVASEAWLTSDLVMGLPGMAQMGTVLGFLQRGA 329
Db        273 VIVVFLRKFHVFNLFTKAIERKIS-KIWIASDNWSTATKIITIPNVKKLGKVVGFAFRRG 331

Qy        330 QLHEFPQYVKT-HL----------ALATDPAFCSALGEREQGLEEDVVGQRCPQC----- 373
Db        332 NTSSFHSFLQTLHMYPNDNNKPLHEFAMLVSACKYIKDGD----------LSQCISNYS 380

Qy        374 -DCITLQNVSAGLNH---------HQT-----FSVYAAVYSVAQALHNTLQCNASGCPAQ 418
Db        381 QATLTYDTTKTIENHLFKRNDFLWHYTEPGLIYSIQLAVFALGHAIRDL--CQARDCKKP 438

Qy        419 DPVKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRF---- 474
Db        439 NAFQPWELLAVLKNVTFTDGRNSFHFDAHGDLNTGYDVVLWKETNGLMTVTKMAEYDLQR 498

Qy        475 ------NGSLRTERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGF-HSCCYDCVDCEAG 527
Db        499 DVFITTNQETKHEFRKLK-------QILSKCSKECSPGQMKKATGSQHSCCYECVSCPEN 551

Qy        528 SYRQNPDDIACTFCGQD-EWSPERSTRCFRRRSRFLAWGEPAVLLLLLLSLALGLVLAA 586
Db        552 HYSNETDMDHCLLCNNETHWAPVRSTTCFEKEVEYLDWDDSLALLLIALSLLGIAFVLAI 611

Qy        587 LGLFVHHRDSPLVQASGGPLACF-GLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLT 645
Db        612 GIIFTRNLKTPVVKSSGGLVVCYVMLICHALNFASTGFFIGEPQDFACKTRQTLFGVSFT 671

Qy        646 GCLSTLFLQAAEIFVESELPLSWADRLS---GCLRGPWAWLVVLLAMLVEVALCTWYLVA 702
Db        672 LCVSCILTKSLKIL----LAFSFDPKLTMFLKCLYRPVP--IVLTCTGIQVVICTLWLVL 725

Qy        703 FPPEVVTDWHMLPTEALVHCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRAR 762
Db        726 AAPSVEEN-ISLPRVIILECEEGSALAFGTMLGYITVLAFICFVFAFKGRKLPENYNEAK 784

Qy        763 GLTFAMLAYFITWVSFVPLLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPG 822
Db        785 FLTFGMLIYFIAWITFIPVYTTTFGKYLPAVEIIVILISNYGILCCIFFPKCYIILCKQK 844

Qy        823 LNTPEFFL 830
```

```
                ||   ||
Db        845 TNTKSAFL 852


RESULT 9
US-10-841-129-6
; Sequence 6, Application US/10841129
; Publication No. US20050250113A1
; GENERAL INFORMATION:
;  APPLICANT: Zuker, Charles S.
;  APPLICANT:  Erlenbach, Isolda
;  APPLICANT:  Hoon, Mark A.
;  APPLICANT:  Ryba, Nicholas J. P.
;  APPLICANT:  Zhang, Yifeng
;  APPLICANT:  The Regents of the University of California
;  APPLICANT:  The Government of the United States of America
;  APPLICANT:  as represented by The Secretary of the
;  APPLICANT:  Department of Health and Human Services
;  TITLE OF INVENTION: A Mammalian Magnesium/Manganese Sensing G Protein
;  TITLE OF INVENTION:  Coupled Receptor
;  FILE REFERENCE: 02307E-145400US
;  CURRENT APPLICATION NUMBER: US/10/841,129
;  CURRENT FILING DATE:  2004-05-07
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
;   LENGTH: 928
;   TYPE: PRT
;   ORGANISM: Rattus sp.
;   FEATURE:
;   OTHER INFORMATION: rat magnesium/manganese sensing G protein coupled
;   OTHER INFORMATION: receptor (GPCR) R5.24, Mg receptor, Mn receptor
US-10-841-129-6

  Query Match             24.0%;  Score 1086.5;  DB 6;  Length 928;
  Best Local Similarity   31.7%;  Pred. No. 2.6e-79;
  Matches  266;  Conservative  158;  Mismatches  353;  Indels   63;  Gaps   20;

Qy         33 GDYVLGGLFPLGEAEEAGLRSRTRPSSP---VCTRFSSNGLLWALAMKMAVEEINNKSDL 89
              |  ::|||| :    |  | |   |  |    |  |  :  |  |||  ::| ||| | |
Db         34 GHIMIGGLFAI---HEKMLSSDDHPRQPQIQKCVGFEISVFLQTLAMIHSIEMINN-SSL 89

Qy         90 LPGLRLGYDLFDTCSEPVVAMKPSLMFLAKAG-SRDIAAY-CNYTQYQPRVLAVIGPHSS 147
              | |::|||:::|||:|   ||  :| ||:|   ||:   : |:|: | ||| |:||   |
Db         90 LSGVKLGYEIYDTCTEVTAAMAATLRFLSKFNCSRETVIFQCDYSSYVPRVKAIIGAGYS 149

Qy        148 ELAMVTGKFFSFFLMPQVSYGASMELLSARETFPSFFRTVPSDRVQLTAAAELLQEFGWN 207
              |::|   :  :  ||||||| :: |:|| :  |||| ||||||  |  | | |:::  |||
Db        150 EISMAVSRMLNLQLMPQVSYESTAEILSDKIRFPSFLRTVPSDFYQTKAMAHLIRQSGWN 209

Qy        208 WVAALGSDDEYGRQGLSIFSALAAARGICIAHEGLVPLPRADDSRLGKVQDVLHQ-VNQS 266
              || |: :||:|||  |: |:  ||   :||| : ::|   :|::   ::   | : : ::
Db        210 WVGAITTDDDYGRLALNTFAIQAAENNVCIAFKEVLPAFLSDNTIEVRINQTLEKIIAEA 269

Qy        267 SVQVVLLFASVHAAHALFNYSISSRLSPKVWVASEAWLTSDLVMGLPGMAQMGTVLGFLQ 326
              | |:::|         ||| :|  ::| |:|:||: | |:  :: :| : ::| |:||
Db        270 QVNVIVVFLRKFHVFNLFNKAIERKIS-KIWIASDNWSTAAKIITIPNVKKLGKVVGFTF 328

Qy        327 RGAQLHEFPQYVKTHLALATD---------PAFCSALGEREQGLEEDVVGQRCPQCDCIT 377
              |   :  |  :::|      :|           | :    ::  | : :             |
Db        329 RRGNMSSFHSFLQTLHMYPSDNNKPLHEFAMLFSACKHIKDGDLSQCISNYSQATWTYDT 388

Qy        378 LQNVSAGLNHHQTF----------SVYAAVYSVAQALHNTLQCNASGCPAQDPVKPWQL 426
               : :   |     |           |: || ::  |: :   |     |   :  :||:|
Db        389 TKTIETHLFKRNDFLWHYTEPGLIHSIQLAVLALGHAIRDL--CQDRDCQKPNAFQPWEL 446

Qy        427 LENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRL---------HDVG-RFNG 476
              |  : |:||  |     ||: |:::  |:: ||     :  :           |||    |
Db        447 LAVLKNVTFTDGKNSFHFDAHGDLNTGYEVVLWKETNGLMTVTKMAEYDLQHDVFITTNQ 506

Qy        477 SLRTERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGF-HSCCYDCVDCEAGSYRQNPDD 535
              :  |  |::     : :|:||::|  ||:::  |  |||||:||:|    |    |
Db        507 ETKHEFRKLK-------QILSKCSKECIPGQMKKATGSQHSCCYECVNCPENHYSNETDM 559

Qy        536 IACTFCGQD-EWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHR 594
                |  |  :  |:| ||| || :   :| | :   |||: |  | :  |||   :|  :
Db        560 DHCLVCNNETHWAPVRSTMCFEKEVEYLDWDDSLALLLIALSLLGIAFVLAVGIIFTRNL 619

Qy        595 DSPLVQASGGPLACF-GLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFL 653
               :|:|::||| : |:  | |  |    |   | |:|    |  :| |  :  | |:| :
Db        620 KTPVVKSSGGLVVCYVMLACHALNFASTGFFIGEPQDFTCKTRQTLFGVSFTLCVSCILT 679

Qy        654 QAAEIFVESELPLSWADRLS---GCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTD 710
              :: :|     | |:  |     || |    :||    ::| :|| :||   |  | :
Db        680 KSLKIL----LAFSFDPTLKTFLKCLYRPVP--IVLTCTGIQVVICTLWLVLAAP-TVEE 732

Qy        711 WHMLPTEALVHCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLA 770
                ||   :: |   | ::||     | |||:||:  |  |  |  || |: ||| ||
Db        733 NTSLPRVIILECEEGSALAFGTMLGYIAVLAFICFVFAFKGRKLPENYNEAKFLTFGMLI 792

Qy        771 YFITWVSFVPLLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFL 830
              ||| |::|:|: |       |||::  :|:   |||     |:||::: :   ||   ||
Db        793 YFIAWITFIPVYATTFGKYLPAVEIIVILISNYGILCCTFFPKCYIILCKQKTNTKSVFL 852


RESULT 10
US-10-841-129-2
; Sequence 2, Application US/10841129
; Publication No. US20050250113A1
; GENERAL INFORMATION:
;  APPLICANT: Zuker, Charles S.
;  APPLICANT:  Erlenbach, Isolda
;  APPLICANT:  Hoon, Mark A.
;  APPLICANT:  Ryba, Nicholas J. P.
;  APPLICANT:  Zhang, Yifeng
;  APPLICANT:  The Regents of the University of California
;  APPLICANT:  The Government of the United States of America
;  APPLICANT:  as represented by The Secretary of the
;  APPLICANT:  Department of Health and Human Services
;  TITLE OF INVENTION: A Mammalian Magnesium/Manganese Sensing G Protein
;  TITLE OF INVENTION:  Coupled Receptor
;  FILE REFERENCE: 02307E-145400US
;  CURRENT APPLICATION NUMBER: US/10/841,129
;  CURRENT FILING DATE:  2004-05-07
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 926
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: human magnesium/manganese sensing G protein
;   OTHER INFORMATION: coupled receptor (GPCR) R5.24, Mg receptor, Mn
;   OTHER INFORMATION: receptor
US-10-841-129-2

  Query Match             23.5%;  Score 1061.5;  DB 6;  Length 926;
  Best Local Similarity   30.3%;  Pred. No. 2.6e-77;
  Matches  260;  Conservative  167;  Mismatches  331;  Indels   99;  Gaps   23;

Qy         33 GDYVLGGLFPLGEAEEAGLRSRTRPSSPVCTRFSSNGLLWALAMKMAVEEINNKSDLLPG 92
              |  ::|||| : |   :   |  ||    |  |  :  |  |||  ::| ||| | ||||
Db         34 GHIIIGGLFAIHEKMLSSEDSPRRPQIQECVGFEISVFLQTLAMIHSIEMINN-STLLPG 92

Qy         93 LRLGYDLFDTCSEPVVAMKPSLMFLAKAG-SRDIAAY-CNYTQYQPRVLAVIGPHSSELA 150
              ::|||:::|||:|   |||  :| ||:|   ||:   : |:|: | ||| ||||    ||:
Db         93 VKLGYEIYDTCTEVTVAMAATLRFLSKFNCSRETVEFKCDYSSYMPRVKAVIGSGYSEIT 152

Qy        151 MVTGKFFSFFLMPQVSYGASMELLSARETFPSFFRTVPSDRVQLTAAAELLQEFGWNWVA 210
              |  :  :  ||||| | :: |:|| :  |||| ||||||  |: | | |:|: ||||:
Db        153 MAVSRMLNLQLMPQVGYESTAEILSDKIRFPSFLRTVPSDFHQIKAMAHLIQKSGWNWIG 212

Qy        211 ALGSDDEYGRQGLSIFSALAAARGICIAHEGLVPLPRADDSRLGKVQDVLHQ-VNQSSVQ 269
              : :||:|||  |: |   | |  :||| : ::|   :|::   ::   | : : :: |
Db        213 IITTDDDYGRLALNTFIIQAEANNVCIAFKEVLPAFLSDNTIEVRINRTLKKIILEAQVN 272
```



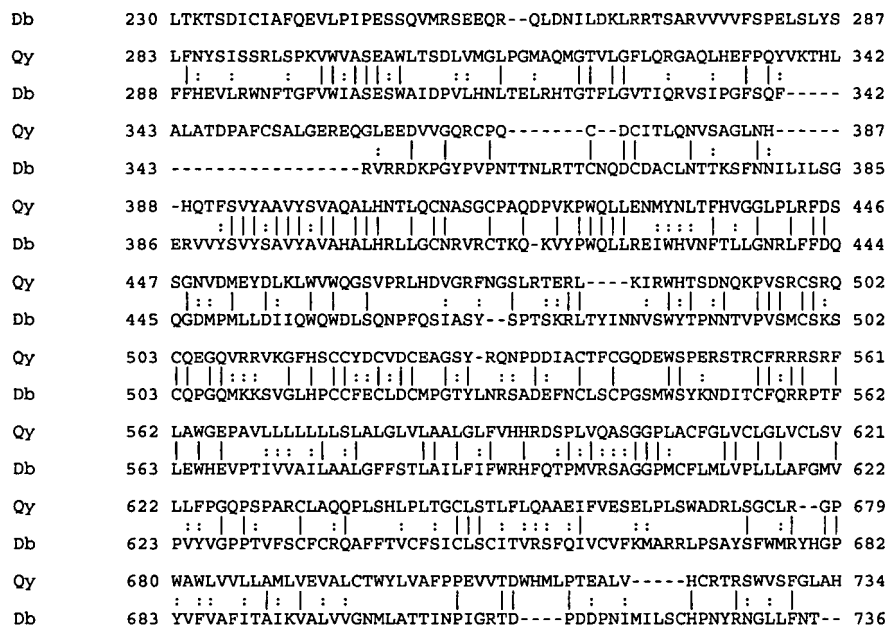
RESULT 11
US-10-645-441-3
; Sequence 3, Application US/10645441
; Publication No. US20050260599A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Ryba, Nicholas J.P.
; APPLICANT: Nelson, Greg
; APPLICANT: Hoon, Mark A.
; APPLICANT: Chandrashekar, Jayaram
; APPLICANT: Zhang, Yifeng
; APPLICANT: The Regents of the University of California
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by the Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: Mammalian Sweet Taste Receptors
; FILE REFERENCE: 02307E-120110US
; CURRENT APPLICATION NUMBER: US/10/645,441
; CURRENT FILING DATE: 2003-08-20
; PRIOR APPLICATION NUMBER: US/09/927,315
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: US 60/302,898
; PRIOR FILING DATE: 2001-07-03
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human T1R1 sweet taste receptor
US-10-645-441-3

Query Match 23.3%; Score 1053; DB 6; Length 777;
Best Local Similarity 32.4%; Pred. No. 9.9e-77;
Matches 255; Conservative 144; Mismatches 339; Indels 48; Gaps 15;



RESULT 12
US-10-645-441-9
; Sequence 9, Application US/10645441
; Publication No. US20050260599A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Ryba, Nicholas J.P.

```
;  APPLICANT:  Nelson, Greg
;  APPLICANT:  Hoon, Mark A.
;  APPLICANT:  Chandrashekar, Jayaram
;  APPLICANT:  Zhang, Yifeng
;  APPLICANT:  The Regents of the University of California
;  APPLICANT:  The Government of the United States of America
;  APPLICANT:  as represented by the Secretary of the
;  APPLICANT:  Department of Health and Human Services
;  TITLE OF INVENTION: Mammalian Sweet Taste Receptors
;  FILE REFERENCE: 02307E-120110US
;  CURRENT APPLICATION NUMBER: US/10/645,441
;  CURRENT FILING DATE:  2003-08-20
;  PRIOR APPLICATION NUMBER: US/09/927,315
;  PRIOR FILING DATE: 2001-08-10
;  PRIOR APPLICATION NUMBER: US 60/302,898
;  PRIOR FILING DATE: 2001-07-03
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
;   LENGTH: 838
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: human T1R2 sweet taste receptor
US-10-645-441-9

  Query Match             22.3%;  Score 1011;  DB 6;  Length 838;
  Best Local Similarity   30.7%;  Pred. No. 2.5e-73;
  Matches  265;  Conservative  153;  Mismatches  372;  Indels   74;  Gaps   23;

Qy          2 LGPAV-----LGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTR 56
              :||        |   || |   |   :   |        |||:|||||  |   |    |:
Db          1 MGPRAKTICSLFFLLWVLAEPAENSDFYL------PGDYLLGGLFSL-HANMKGIVHLNF 53

Qy         57 PSSPVCTRFSSNGLLWAL--AMKMAVEEINNKSDLLPGLRLGYDLFDTC--SEPVVAMKP 112
                |:|  :    : : |  ||: ||||||| | ||||: |||:: | |  |  |   :|
Db         54 LQVPMCKEYEVKVIGYNLMQAMRFAVEEINNDSSLLPGVLLGYEIVDVCYISNNV---QP 110

Qy        113 SLMFLAKAGSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASME 172
              | |||   :  :    :|: |  ||:||||| :||  |    | | ||:||::| |  :
Db        111 VLYFLAHEDNL-LPIQEDYSNYISRVVAVIGPDNSESVMTVANFLSLFLLPQITYSAISD 169

Qy        173 LLSARETFPSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAA 232
              |  :  ||:  || ||    : |  :|:  | |||:  | | | |||    :      |
Db        170 ELRDKVRFPALLRTTPSADHHVEAMVQLMLHFRWNWIIVLVSSDTYGRDNGQLLGERVAR 229

Qy        233 RGICIAHEGLVPLPRADDSRLG----KVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSI 288
              | |||| :  :|  : : :       ::  :: :: ||: :||::|:    :  ||  :
Db        230 RDICIAFQETLPTLQPNQNMTSEERQRLVTIVDKLQQSTARVVVVFSPDLTLYHFFNEVL 289

Qy        289 SSRLSPKVWVASEAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEF-------PQYVKTH 341
                :  ||:|||:|    ::  |  :  :|| ||   :   :  |        ||
Db        290 RQNFTGAVWIASESWAIDPVLHNLTELGHLGTFLGITIQSVPIPGFSEFREWGPQAGPPP 349

Qy        342 LALATDPAFCSALGEREQGLEEDVVGQRCPQCDCITLQ-NVSAGLNHHQ-TFSVYAAVYS 399
              |:  :    |:              | |  |   ||  |    |:  :  :|||:|||:
Db        350 LSRTSQSYTCN--------------QECDNCLNATLSFNTILRLSGERVVYSVYSAVYA 394

Qy        400 VAQALHNTLQCNASGCPAQDPVKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLW 459
              || |||: | |: | |  :  | |||||| :: : | :    : ||  |:| :  ::  |
Db        395 VAHALHSLLGCDKSTC-TKRVVYPWQLLEEIWKVNFTLLDHQIFFDPQGDVALHLEIVQW 453

Qy        460 VWQGSVPRLHDVGRFNGSLRTERLK----IRWHTSDNQKPVSRCSRQCQEGQVRRVKGFH 515
              |  |       |  :    |  :||    | ||| :|  |:| ||::|| || ::  | |
Db        454 QWDRSQNPFQSVASYYPLQR--QLKNIQDISWHTVNNTIPMSMCSKRCQSGQKKKPVGIH 511

Qy        516 SCCYDCVDCEAGSY-RQNPDDIACTFCGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLL 574
              ||::|:||  |::     |:  |  |  :||| :  | ||:|:  || | |   : : |
Db        512 VCCFECIDCLPGTFLNHTEDEYECQACPNNEWSYQSETSCFKRQLVFLEWHEAPTIAVAL 571

Qy        575 LLSLALGLVLAALGLFVHHRDSPLVQASGGPLACFGLVCLGLVCLSVL-LFPGQPSPARC 633
              | :|     || | :|  |  :|:|:::|||: || :: | ||   |: :: | |  : |
Db        572 LAALGFLSTLAILVIFWRHFQTPIVRSAGGPM-CFLMLTLLLVAYMVVPVYVGPPKVSTC 630

Qy        634 LAQQPLSHLPLTGCLSTLFLQAAEIFVESELPLSWADRLSGCLR--GPW---AWLVVLLA 688
              | :| |  |  | |:| : ::: :|    ::   :   |  :|  ||:   |:: ||
Db        631 LCRQALFPLCFTICISCIAVRSFQIVCAFKMASRFPRAYSYWVRYQGPYVSMAFITVLKM 690

Qy        689 MLVEVALCTWYLVAFP---PEVVTDWHMLPTEALVHCRTRSWVSFGLAHATNATLAFLCF 745
              . ::| :      ::| |   |     |     |   :| |       |      : :  |: : |
Db        691 VIVVIG-----MLARPQSHPRTDPD---DPKITIVSCNPNYRNSLLFNTSLDLLLSVVGF 742

Qy        746 LGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVPLLANVQVVLRPAVQMGALLLCVLGI 805
                :: :  |  || |: :| :|  || : ||    ::    ||   | :   :| :| |
Db        743 SFAYMGKELPTNYNEAKFITLSMTFYFTSSVSLCTFMSAYSGVLVTIVDLLVTVLNLLAI 802

Qy        806 LAAFHLPRCYLLMRQPGLNTPEFF 829
                :  |:||:::  |  ||| :|
Db        803 SLGYFGPKCYMILFYPERNTPAYF 826


RESULT 13
US-10-645-441-7
; Sequence 7, Application US/10645441
; Publication No. US20050260599A1
; GENERAL INFORMATION:
;  APPLICANT: Zuker, Charles S.
;  APPLICANT:  Ryba, Nicholas J.P.
;  APPLICANT:  Nelson, Greg
;  APPLICANT:  Hoon, Mark A.
;  APPLICANT:  Chandrashekar, Jayaram
;  APPLICANT:  Zhang, Yifeng
;  APPLICANT:  The Regents of the University of California
;  APPLICANT:  The Government of the United States of America
;  APPLICANT:  as represented by the Secretary of the
;  APPLICANT:  Department of Health and Human Services
;  TITLE OF INVENTION: Mammalian Sweet Taste Receptors
;  FILE REFERENCE: 02307E-120110US
;  CURRENT APPLICATION NUMBER: US/10/645,441
;  CURRENT FILING DATE:  2003-08-20
;  PRIOR APPLICATION NUMBER: US/09/927,315
;  PRIOR FILING DATE: 2001-08-10
;  PRIOR APPLICATION NUMBER: US 60/302,898
;  PRIOR FILING DATE: 2001-07-03
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
;   LENGTH: 843
;   TYPE: PRT
;   ORGANISM: Rattus sp.
;   FEATURE:
;   OTHER INFORMATION: rat T1R2 sweet taste receptor
US-10-645-441-7

  Query Match             22.2%;  Score 1002.5;  DB 6;  Length 843;
  Best Local Similarity   30.2%;  Pred. No. 1.2e-72;
  Matches  265;  Conservative  159;  Mismatches  357;  Indels   97;  Gaps   25;

Qy          2 LGP-----AVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTR 56
              :||       :| | |  |  ||    |  :    : |||:||||| |     | ::| :
Db          1 MGPQARTLCLLSLLLHVLPKPG---KLVENSDFHLAGDYLLGGLFTL----HANVKSISH 53

Qy         57 PS---SPVCTRFSSNGLLWAL--AMKMAVEEINNKSDLLPGLRLGYDLFDTC--SEPVVA 109
                |    | |  |:   | : |  ||: ||||||| | ||||: |||:: | |  |  :
Db         54 LSYLQVPKCNEFTMKVLGYNLMQAMRFAVEEINNCSSLLPGVLLGYEMVDVCYLSNNI-- 111

Qy        110 MKPSLMFLAKAGSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGA 169
                | | |||:     :    :|:|| | |:||||| :|| |:       | ||:||::| |
Db        112 -HPGLYFLAQDDDL-LPILKDYSQYMPHVVAVIGPDNSESAITVSNILSHFLIPQITYSA 169

Qy        170 SMELLSARETFPSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFS-A 228
              : |  :  |||  |||||    : |  :|:  | |||:  | |||:|||:   : |
Db        170 ISDKLRDKRHFPSMLRTVPSATHHIEAMVQLMVHFQWNWIVVLVSDDDYGRENSHLLSQR 229

Qy        229 LAAARGICIAHEGLVPLP------RADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHA 282
              |     |||| : ::|:|      |::: |  :: ::| :: ::| :||::|:   : ::
```



RESULT 14
US-10-645-441-8
; Sequence 8, Application US/10645441
; Publication No. US20050260599A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Ryba, Nicholas J.P.
; APPLICANT: Nelson, Greg
; APPLICANT: Hoon, Mark A.
; APPLICANT: Chandrashekar, Jayaram
; APPLICANT: Zhang, Yifeng
; APPLICANT: The Regents of the University of California
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by the Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: Mammalian Sweet Taste Receptors
; FILE REFERENCE: 02307E-120110US
; CURRENT APPLICATION NUMBER: US/10/645,441
; CURRENT FILING DATE: 2003-08-20
; PRIOR APPLICATION NUMBER: US/09/927,315
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: US 60/302,898
; PRIOR FILING DATE: 2001-07-03
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 843
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: mouse T1R2 sweet taste receptor
US-10-645-441-8

Query Match 22.1%; Score 999.5; DB 6; Length 843;
Best Local Similarity 30.5%; Pred. No. 2.1e-72;
Matches 268; Conservative 157; Mismatches 355; Indels 99; Gaps 26;

RESULT 15
US-10-821-234-1133
; Sequence 1133, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:

```
;  APPLICANT: Labat, Ivan
;  APPLICANT:  Stache-Crain, Birgit
;  APPLICANT:  Andarmani, Susan
;  APPLICANT:  Tang, Y. Tom
;  TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
;  FILE REFERENCE: 821A
;  CURRENT APPLICATION NUMBER: US/10/821,234
;  CURRENT FILING DATE:  2004-04-07
;  PRIOR APPLICATION NUMBER: US 60/462,047
;  PRIOR FILING DATE: 2003-04-07
;  NUMBER OF SEQ ID NOS: 1704
;  SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 1133
;   LENGTH: 2919
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-821-234-1133

  Query Match             2.1%;  Score 96.5;  DB 6;  Length 2919;
  Best Local Similarity   19.4%;  Pred. No. 19;
  Matches  211;  Conservative  105;  Mismatches  341;  Indels  431;  Gaps   59;

Qy          3 GPAVLGLSLWALLHPGTGAPLCLS-----------------QQL----RMKGDYVLGGLF 41
              |  :: :: |||  |  | |  |                  |||    |: |  |    |
Db          9 GGWLVDVTAWALSPPAVGTPQPLGYITARQSGAVADLIEPLQQLRAWPRVGGTIVQQLSF 68

Qy         42 P-------------------------LGEAEEAGLRSRTRPSSPVCTRFSSNGLLWALAM 76
              |                         ||::    |: :  |:     ||:  |:| |||:
Db         69 PGTLEIAAALICRGRHHCPAAEFPRDLGDSRSPHLQGKMIPA-----RFA--GVLLALAL 121

Qy         77 ----KMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKAGSR-DIAAYCNY 131
                  :  |    :|       | |  :|                  ||    | ||:|
Db        122 ILPGTLCAEGTRGRSSTARCSLFGSDFVNTFD---------------GSMYSFAGYCSY 165

Qy        132 T---QYQPRVLAVIGPHSS----ELAMVTGKFFSFFL--------------MPQVSYGAS 170
                  | |  ::||   :     |::  |:||   |                ||  | |
Db        166 LLAGGCQKRSFSIIGDFQNGKRVSLSVYLGEFFDIHLFVNGTVTQGDQRVSMPYASKGLY 225

Qy        171 MEL-------------LSARETFPSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDE 217
              :|                ||       |: : |||         |   |       : ::|:
Db        226 LETEAGYYKLSGEAYGFVARIDGSGNFQVLLSDR-YFNKTCGLCGNFN------IFAEDD 278

Qy        218 YGRQGLSIFS---------ALAAARGICIAHEGLVPLPRADDSRLGKVQDVLHQVNQSSV 268
              :  |  :: |          ||::    |   |    |  : :   |::|  | :  |
Db        279 FMTQEGTLTSDPYDFANSWALSSGEQWC---ERASPPSSSCNISSGEMQKGLWEQCQ--- 332

Qy        269 QVVLLFASVHA-AHALFNYSISSRLSPKVWVASEAWLTSDLVMGL----PGM-------A 316
                :|   || |  | |         : |: :||        :   ||     | :       |
Db        333 --LLKSTSVFARCHPL--------VDPEPFVALCEKTLCECAGGLECACPALLEYARTCA 382

Qy        317 QMGTVL-GFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQGLEEDVVGQRCP---- 371
              | | || |:     |       |  ::    ::      | :|    |   |  | |  ||
Db        383 QEGMVLYGWTDHSACSPVCPAGMEYRQCVSPCARTCQSLHINEMCQERCVDGCSCPEGQL 442

Qy        372 ----------QCDCITL---------------------------------QNVSAGLNHH 388
                       :| |:                                   : :  | :|
Db        443 LDEGLCVESTECPCVHSGKRYPPGTSLSRDCNTCICRNSQWICSNEECPGECLVTGQSHF 502

Qy        389 QTF-SVYAAVYSVAQAL-------HN------TLQC---NASGCPAQDPVKPWQLLENMY 431
              ::| : |      : | |        |:      |:||     : |     |:    |  ::
Db        503 KSFDNRYFTFSGICQYLLARDCQDHSFSIVIETVQCADDRDAVCTRSVTVR----LPGLH 558

Qy        432 N-------------------LTFHVGGLPLRFDSSGNVDMEY--DLKL-WVWQGS-VPRL 468
              |                   |    | | ::   : :| : |  ||:: |  :|  : :|
Db        559 NSLVKLKHGAGVAMDGQDVQLPLLKGDLRIQHTVTASVRLSYGEDLQMDWDGRGRLLVKL 618

Qy        469 HDV---------GRFNGSLRTERL-----------------KIRWHTSDNQKPVS----- 497
                |         | :||:   : |                 |:     | ||  |
Db        619 SPVYAGKTCGLCGNYNGNQGDDFLTPSGLAEPRVEDFGNAWKLHGDCQDLQKQHSDPCAL 678

Qy        498 --RCSRQCQEG--------------QVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTFC 541
                | :|  :|                |  :    :| ||   |  |          |  |
Db        679 NPRMTRFSEEACAVLTSPTFEACHRAVSPLPYLRNCRYDVCSCSDGR--------EC-LC 729

Qy        542 GQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQA 601
              |       :  |  |  | :|| ||    |                     : | |
Db        730 GA---LASYAAACAGRGVR-VAWREPGRCEL---------------------NCPKGQV 763

Qy        602 SGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFVE 661
                  | | |  |  | | | | :| :     |          | ||
Db        764 Y---LQC-GTPC-NLTCRS-LSYPDEECNEAC----------LEGCFCP----------- 796

Qy        662 SELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYL---VAFPPEVVTDWHML--PT 716
                |  : |    |                 |  | |  |    :  | :: :| | :
Db        797 ---PGLYMDERGDC--------------VPKAQCPCYYDGEIFQPEDIFSDHHTMCYCE 838

Qy        717 EALVHCRTRSWV----------SFGLAHATNATLA-------FLCFLGTFLVRSQPGCYN 759
              :  :|| | | |          |  |:| :  :|:       :|          |
Db        839 DGFMHC-TMSGVPGSLLPDAVLSSPLSHRSKRSLSCRPPMVKLVC----------PADNL 887

Qy        760 RARGL--TFAMLAYFITWVSFVPLLANVQVVLRPAVQMGALLLCVLGILA----AFHLPR 813
              || ||  |      | :  :|                 ||    || | |        | |
Db        888 RAEGLECTKTCQNYDLECMS----------------MG----CVSGCLCPPGMVRHENR 926

Qy        814 CYLLMRQP 821
              | | | |
Db        927 CVALERCP 934


Search completed: December 16, 2005, 18:15:35
Job time : 17 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2005, 17:55:56 ; Search time 190 Seconds
(without alignments)
1970.267 Million cell updates/sec

Title: US-10-679-102-15
Perfect score: 4524
Sequence: 1 MLGPAVLGLSLWALLHPGTG.....GPGDAQGQNDGNTGNQGKHE 852

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 4524 | 100.0 | 852 | 5 | ABB79827 | Abb79827 Human tas |
| 2 | 4524 | 100.0 | 852 | 7 | ADK90725 | Adk90725 Human tas |
| 3 | 4524 | 100.0 | 852 | 9 | ADZ36810 | Adz36810 Human T1R |
| 4 | 4524 | 100.0 | 936 | 7 | ADC86149 | Adc86149 Human GPC |
| 5 | 4517 | 99.8 | 852 | 7 | ADC26323 | Adc26323 Human NOV |
| 6 | 4512 | 99.7 | 852 | 4 | AAU08996 | Aau08996 Human G p |
| 7 | 4512 | 99.7 | 852 | 5 | AAO21501 | Aao21501 Human G-p |
| 8 | 4512 | 99.7 | 852 | 5 | AAU78587 | Aau78587 Human AXO |
| 9 | 4512 | 99.7 | 852 | 6 | ABP70451 | Abp70451 Amino aci |
| 10 | 4512 | 99.7 | 852 | 6 | ABU62106 | Abu62106 Human cal |
| 11 | 4512 | 99.7 | 852 | 8 | ADG19752 | Adg19752 Human G p |
| 12 | 4512 | 99.7 | 852 | 8 | ADG73581 | Adg73581 Human T1R |
| 13 | 4512 | 99.7 | 852 | 8 | ADR29260 | Adr29260 Taste rec |
| 14 | 4512 | 99.7 | 852 | 9 | ADW26968 | Adw26968 Human tas |
| 15 | 4512 | 99.7 | 852 | 9 | ADY27327 | Ady27327 Synthetic |
| 16 | 4509 | 99.7 | 852 | 5 | AAU73184 | Aau73184 Human SAC |
| 17 | 4509 | 99.7 | 852 | 8 | ADM15735 | Adm15735 Human SAC |
| 18 | 4506 | 99.6 | 852 | 6 | ABG71996 | Abg71996 Human tas |
| 19 | 4500 | 99.5 | 850 | 4 | AAE10366 | Aae10366 Human tas |
| 20 | 4496.5 | 99.4 | 863 | 5 | ABB77318 | Abb77318 Human G-p |
| 21 | 4496.5 | 99.4 | 863 | 8 | ADI40942 | Adi40942 Human G p |
| 22 | 4493 | 99.3 | 850 | 7 | ADK90739 | Adk90739 Human tas |
| 23 | 4490 | 99.2 | 850 | 6 | ABR43489 | Abr43489 Human swe |
| 24 | 4474 | 98.9 | 894 | 8 | ADI40946 | Adi40946 Human G p |
| 25 | 4464.5 | 98.7 | 849 | 9 | ADZ36826 | Adz36826 Human T1R |
| 26 | 4461 | 98.6 | 854 | 7 | ADJ93196 | Adj93196 Rat taste |
| 27 | 4461 | 98.6 | 854 | 7 | ADJ93109 | Adj93109 Human G-c |
| 28 | 4461 | 98.6 | 854 | 8 | ADI41026 | Adi41026 Human GPC |
| 29 | 4365.5 | 96.5 | 829 | 8 | ADI40944 | Adi40944 Human G p |
| 30 | 4222 | 93.3 | 805 | 7 | ADC12740 | Adc12740 Human GPC |
| 31 | 4161.5 | 92.0 | 807 | 5 | AAO15707 | Aao15707 Human pro |
| 32 | 4151.5 | 91.8 | 853 | 9 | ADY27344 | Ady27344 Synthetic |
| 33 | 4145.5 | 91.6 | 855 | 9 | ADY27338 | Ady27338 Synthetic |
| 34 | 4002 | 88.5 | 771 | 6 | ABP81709 | Abp81709 Human G p |
| 35 | 3708 | 82.0 | 855 | 5 | AAU80493 | Aau80493 Human G-c |
| 36 | 3605.5 | 79.7 | 857 | 9 | ADY27345 | Ady27345 Synthetic |
| 37 | 3599.5 | 79.6 | 859 | 9 | ADY27339 | Ady27339 Synthetic |
| 38 | 3371 | 74.5 | 865 | 9 | ADW26958 | Adw26958 Feline ta |
| 39 | 3255 | 71.9 | 858 | 4 | AAE10371 | Aae10371 Rat taste |
| 40 | 3255 | 71.9 | 858 | 5 | ABB79832 | Abb79832 Rat taste |
| 41 | 3255 | 71.9 | 858 | 6 | ABR43495 | Abr43495 Rat sweet |
| 42 | 3255 | 71.9 | 858 | 6 | ABP70448 | Abp70448 Amino aci |
| 43 | 3255 | 71.9 | 858 | 7 | ADK90735 | Adk90735 Rat taste |
| 44 | 3255 | 71.9 | 858 | 8 | ADG73578 | Adg73578 Rat T1R3. |
| 45 | 3255 | 71.9 | 858 | 8 | ADR29257 | Adr29257 Taste rec |

ALIGNMENTS

RESULT 1
ABB79827
ID ABB79827 standard; protein; 852 AA.
XX
AC ABB79827;
XX
DT 25-NOV-2002 (first entry)
XX
DE Human taste cell specific receptor hT1R3.
XX
KW hT1R3; T1R3; T1R; G-protein coupled receptor; receptor; human; taste;
KW sweet.
XX
OS Homo sapiens.
XX
PN WO200264631-A2.
XX
PD 22-AUG-2002.
XX
PF 03-JAN-2002; 2002WO-US000198.
XX
PR 03-JAN-2001; 2001US-0259227P.
PR 19-APR-2001; 2001US-0284547P.
XX
PA (SENO-) SENOMYX INC.
XX
PI Adler JE, Li X, Staszewski L, O'connell S, Zozulya S;
XX
DR WPI; 2002-666991/71.
DR N-PSDB; ABN84919, ABN84920.
XX
PT Novel mammalian taste-cell specific G protein-coupled receptors active in
PT taste signaling, useful for representing the perception of taste and for
PT predicting the perception of taste in a mammal.
XX
PS Claim 152; Page 76; 132pp; English.
XX
CC The present sequence is the protein sequence of a novel human G-protein
CC coupled receptor (GPCR), designated hT1R3, that is involved in sweet
CC taste transduction. The invention provides a new family of taste cell-
CC specific GPCRs, the T1Rs, active in taste perception. Genomic DNAs and
CC cDNAs encoding the receptors are described, along with methods for
CC isolating T1R genes and for isolating and expressing T1R polypeptides.
CC Methods for representing taste perception of a particular taste stimulus
CC in a mammal, including a human, are also described, as are methods for
CC generating novel molecules or combinations of molecules that elicit a
CC predetermined taste perception in a mammal, and methods for stimulating

```
CC   one or more tastes. Methods for stimulating or blocking taste perception
CC   are also disclosed, as are methods of screening for modulators of these
CC   novel taste cell-specific GPCRs. Such modulators are useful for
CC   pharmacological, chemical and genetic modulation of taste signalling
CC   pathways. They can be used in the food or pharmaceutical industries to
CC   customize taste, e.g. to modulate the sweet tastes of foods or drugs. T1R
CC   proteins and polypeptides are useful for identifying taste cells present
CC   in the oral cavity, gastrointestinal epithelium and epiglottis. They may
CC   also serve as tools for the generation of taste topographic maps that
CC   elucidate the relationship between the taste cells of the tongue and
CC   taste sensory neurons leading to taste centres in the brain
XX
SQ   Sequence 852 AA;

  Query Match             100.0%;  Score 4524;  DB 5;  Length 852;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  852;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360

Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420

Qy        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540

Qy        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600

Qy        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720

Qy        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780

Qy        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840

Qy        841 NDGNTGNQGKHE 852
              ||||||||||||
Db        841 NDGNTGNQGKHE 852


RESULT 2
ADK90725
ID   ADK90725 standard; protein; 852 AA.
XX
AC   ADK90725;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Human taste receptor T1R3 protein.
XX
KW   gene therapy; taste receptor; T1R3; T1R1; T1R2; taste-induced behavior;
KW   taste topographic map; tongue; taste cell; taste sensory neuron;
KW   brain taste center.
XX
OS   Homo sapiens.
XX
PN   WO2003004992-A2.
XX
PD   16-JAN-2003.
XX
PF   03-JUL-2002; 2002WO-US021269.
XX
PR   03-JUL-2001; 2001US-0302898P.
PR   10-AUG-2001; 2001US-00927315.
PR   22-FEB-2002; 2002US-0358925P.
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Zuker CS,  Ryba NJP,  Chandrashekar J,  Hoon MA,  Nelson GA,  Zhang Y;
XX
DR   WPI; 2003-210382/20.
DR   N-PSDB; ADK90724.
XX
PT   New taste receptors comprising a T1R3 polypeptide and optionally a T1R1
PT   or T1R2 polypeptide, useful as probes for identifying taste cells, and in
PT   gene therapy for correcting acquired and inherited genetic defects.
XX
PS   Claim 3; SEQ ID NO 15; 119pp; English.
XX
CC   The invention relates to a novel taste receptor comprising a T1R3
CC   polypeptide and optionally a T1R1 or T1R2 polypeptide. The nucleic acids
CC   and proteins encoding the receptors may be used as probes to identify
CC   taste cells or subsets of taste cells or to dissect taste-induced
CC   behaviors, and in the generation of taste topographic maps that elucidate
CC   the relationship between the taste cells of the tongue and taste sensory
CC   neurons leading to taste centers in the brain. Nucleic acids encoding T1R
CC   protein can be used with highly density oligonucleotide array technology
CC   to identify T1R protein, orthologs, alleles, conservatively modified
CC   variants, and polymorphic variants in this invention. The nucleic acids
CC   may be inserted into vectors for gene therapy correct acquired and
CC   inherited genetic defects and other diseases. Compositions may be
CC   administered to elicit a therapeutic response in the patient. Modulators
CC   of sweet and/or amino acid taste transduction are useful for
CC   pharmacological and genetic modulation of sweet and amino acid taste
CC   ligands, and in food and pharmaceutical industries to customize taste..
CC   This sequence corresponds to the human T1R3 protein.
XX
SQ   Sequence 852 AA;

  Query Match             100.0%;  Score 4524;  DB 7;  Length 852;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  852;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
```

```
Qy          61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120

Qy         121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180

Qy         181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240

Qy         241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy         301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360

Qy         361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420

Qy         421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy         481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540

Qy         541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600

Qy         601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy         661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720

Qy         721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780

Qy         781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840

Qy         841 NDGNTGNQGKHE 852
               ||||||||||||
Db         841 NDGNTGNQGKHE 852
```

```
RESULT 3
ADZ36810
ID   ADZ36810 standard; protein; 852 AA.
XX
AC   ADZ36810;
XX
DT   30-JUN-2005  (first entry)
XX
DE   Human T1R3 taste receptor protein, SEQ ID NO:15.
XX
KW   Taste receptor; G-protein coupled receptor; GPCR; food; flavour enhancer;
KW   sweetener; screening; T1R3 protein; receptor.
XX
OS   Homo sapiens.
XX
PN   WO2005033125-A2.
XX
PD   14-APR-2005.
XX
PF   04-OCT-2004; 2004WO-US032678.
XX
PR   02-OCT-2003; 2003US-00679102.
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Erlenbach I,  Ryba NJP,  Zhao G,  Zuker CS;
XX
DR   WPI; 2005-285400/29.
DR   N-PSDB; ADZ36808, ADZ36809.
XX
PT   Isolated homodimeric taste receptor used to screen for naturally
PT   occurring or artificial sweet tasting molecules, includes two
PT   polypeptides encoded by nucleotide sequence that hybridizes to nucleotide
PT   sequence encoding amino acid sequence.
XX
PS   Claim 1; SEQ ID NO 15; 107pp; English.
XX
CC   The present invention relates to mammalian taste receptors, the T1Rs and
CC   T2Rs that have been implicated in sweet, umami and bitter taste
CC   detection. The T2Rs are a family of about 30 taste-specific G-protein
CC   coupled receptors (GPCRs) related to opsins, and clustered in regions of
CC   the genome genetically linked to bitter taste in humans and mice. The
CC   T1Rs are a small family of 3 GPCRs expressed in taste cells of the tongue
CC   and palate epithelium, related to metabotropic glutamate receptors. T1Rs
CC   combine to generate at least 2 heteromeric receptors: T1R1 and T1R3 form
CC   an L-amino acid sensor, which in rodents recognizes most amino acids, and
CC   T1R2 and T1R3 associate to function as a broadly tuned sweet receptor.
CC   The new invention provides an isolated homodimeric taste receptor
CC   comprising two T1R3 polypeptides. Also disclosed is a method of
CC   identifying a compound that modulates taste signal transduction in taste
CC   cells, comprising contacting the compound with a homodimeric taste
CC   receptor comprising two T1R3 polypeptides, and determining the functional
CC   effect of the compound upon the receptor, thus identifying a compound
CC   that modulates taste signal transduction. The T1R3 polypeptide comprises
CC   the amino acid sequences given as (SEQ ID NO:15, 20, 23, 25, or 31). The
CC   T1R3 polypeptide is encoded by a nucleotide sequences (SEQ ID NO:14, 19,
CC   22, 24, or 30) given in the specification. The receptor binds to sweet
CC   taste ligands comprising glucose, fructose, galactose, sucrose, maltose,
CC   or lactose. The receptor has G protein coupled receptor activity. The
CC   invention is used to screen for naturally occurring or artificial sweet
CC   tasting molecules or modulators of sweet taste transduction, e.g. small
CC   organic molecules, amino acids, peptides, carbohydrates, lipids, and
CC   polysaccharides. The modulators of sweet taste transduction are used in
CC   food and pharmaceutical industries. The present sequence is the human
CC   T1R3 protein sequence (see ADZ36810).
XX
SQ   Sequence 852 AA;

  Query Match             100.0%;  Score 4524;  DB 9;  Length 852;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  852;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy           1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60

Qy          61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120

Qy         121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180

Qy         181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
```

```
Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360

Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420

Qy        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540

Qy        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600

Qy        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720

Qy        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780

Qy        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840

Qy        841 NDGNTGNQGKHE 852
              ||||||||||||
Db        841 NDGNTGNQGKHE 852
```

```
RESULT 4
ADC86149
ID   ADC86149 standard; protein; 936 AA.
XX
AC   ADC86149;
XX
DT   01-JAN-2004  (first entry)
XX
DE   Human GPCR protein SEQ ID NO:602.
XX
KW   human; GPCR; guanosine triphosphate-binding protein coupled receptor;
KW   gene therapy.
XX
OS   Homo sapiens.
XX
PN   EP1270724-A2.
XX
PD   02-JAN-2003.
XX
PF   18-JUN-2002; 2002EP-00013517.
XX
PR   18-JUN-2001; 2001JP-00246789.
XX
PA   (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
PA   (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX
PI   Suwa M,  Asai K,  Akiyama Y,  Aburatani H;
XX
DR   WPI; 2003-315783/31.
DR   N-PSDB; ADC86148.
XX
PT   New polynucleotide, useful for preparing a composition for treating a
PT   patient in need of increased or suppressed activity or expression of the
PT   guanosine triphosphate-binding protein coupled receptor.
XX
PS   Claim 2; SEQ ID NO 602; 28pp; English.
XX
CC   The invention relates to a novel polynucleotide encoding a guanosine
CC   triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
CC   the invention may have a use in gene therapy. The polynucleotide and
CC   polypeptide are useful for preparing a composition for treating a patient
CC   in need of increased or suppressed activity or expression of the
CC   guanosine triphosphate-binding protein coupled receptor. The protein
CC   sequences shown in ADC85549-ADC87617 represent GPCR's of the invention.
XX
SQ   Sequence 936 AA;
```

Query Match 100.0%; Score 4524; DB 7; Length 936;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 852; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360

Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420

Qy        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540

Qy        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600

Qy        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720

Qy        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780

Qy        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840

Qy        841 NDGNTGNQGKHE 852
                ||||||||||||
Db        841 NDGNTGNQGKHE 852


RESULT 5
ADC26323
ID   ADC26323 standard; protein; 852 AA.
XX
AC   ADC26323;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Human NOV34e protein.
XX
KW   NOV; cytostatic; metabolic disorder; immune; neurodegenerative;
KW   circulatory; haemopoietic; wasting; cancer; gene therapy; vaccine;
KW   transgenic; human.
XX
OS   Homo sapiens.
XX
PN   WO2003004687-A2.
XX
PD   16-JAN-2003.
XX
PF   03-JUL-2002; 2002WO-US021361.
XX
PR   05-JUL-2001; 2001US-0303046P.
PR   09-JUL-2001; 2001US-0303828P.
PR   09-JUL-2001; 2001US-0304016P.
PR   11-JUL-2001; 2001US-0304502P.
PR   13-JUL-2001; 2001US-0305262P.
PR   16-JUL-2001; 2001US-0305673P.
PR   17-JUL-2001; 2001US-0306085P.
PR   24-JUL-2001; 2001US-0307536P.
PR   27-JUL-2001; 2001US-0308228P.
PR   30-JUL-2001; 2001US-0308877P.
PR   01-AUG-2001; 2001US-0309255P.
PR   17-AUG-2001; 2001US-0313328P.
PR   12-SEP-2001; 2001US-0318711P.
PR   19-SEP-2001; 2001US-0323380P.
PR   21-SEP-2001; 2001US-0323969P.
PR   04-JAN-2002; 2002US-0345022P.
PR   04-JAN-2002; 2002US-0345038P.
PR   28-FEB-2002; 2002US-0361172P.
PR   01-MAR-2002; 2002US-0360814P.
PR   01-MAR-2002; 2002US-0360830P.
PR   01-MAR-2002; 2002US-0361133P.
PR   01-MAR-2002; 2002US-0361147P.
PR   05-MAR-2002; 2002US-0361677P.
PR   02-APR-2002; 2002US-0363637P.
PR   12-APR-2002; 2002US-0372326P.
PR   16-APR-2002; 2002US-0372990P.
PR   19-APR-2002; 2002US-0373881P.
PR   19-APR-2002; 2002US-0373921P.
PR   02-JUL-2002; 2002US-00188186.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Anderson DW,  Berghs C,  Boldog FL,  Burgess CE,  Casman SJ;
PI   Catterton E,  Edinger S,  Eisen AJ,  Ellerman K,  Gerlach V,  Gorman L;
PI   Guo X,  Jeffers M,  Kekuda R,  Li L,  Malyankar UM,  Miller CE;
PI   Padigaru M,  Patturajan M,  Pena CEA,  Rastelli L,  Shenoy S;
PI   Shimkets RA,  Spaderna SK,  Spytek KA,  Stone DJ,  Taupier RJ;
PI   Vernet CAM,  Voss EZ,  Zhong M;
XX
DR   WPI; 2003-221607/21.
DR   N-PSDB; ADC26322.
XX
PT   New isolated NOVX polypeptide, useful for determining the presence of, or
PT   predisposition to a disease associated with altered levels of expression
PT   of the polypeptide, and for treating or preventing cancer.
XX
PS   Claim 2; SEQ ID NO 148; 478pp; English.
XX
CC   The invention relates to a novel isolated NOV polypeptide. The
CC   polypeptide of the invention demonstrates cytostatic activity and may be
CC   used for determining the presence of, or predisposition to a disease
CC   associated with altered levels of expression of the polypeptide,
CC   including metabolic disorders, immune disorders, neurodegenerative
CC   disorders, circulatory diseases, haemopoietic disorders, wasting diseases
CC   and cancer. The polypeptide may also be utilised during gene therapy
CC   procedures, vaccine development and transgenic animal production. The
CC   current sequence is that of the human NOV protein of the invention.
XX
SQ   Sequence 852 AA;

  Query Match             99.8%;  Score 4517;  DB 7;  Length 852;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  851;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360

Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420

Qy        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERLKIRWHTSDNQVPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540

Qy        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600

Qy        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
```

```
Qy        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780

Qy        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840

Qy        841 NDGNTGNQGKHE 852
              ||||||||||||
Db        841 NDGNTGNQGKHE 852
```

RESULT 6
AAU08996
ID AAU08996 standard; protein; 852 AA.
XX
AC AAU08996;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human G protein-coupled receptor, GPCR, 50289.
XX
KW Human; GPCR; G protein-coupled receptor; 50289; cardiant;
KW antiatherosclerotic; analgesic; cytostatic; antianginal;
KW cardiovascular disorder; angiogenesis-related disorder; neural disorder;
KW pain response disorder; inflammatory disorder; atherosclerosis;
KW angina pectoris; myocardial infarction; ischaemic heart disease;
KW sudden cardiac death; obesity; hypertensive heart disease; diabetes;
KW prostate cancer-related pain.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 1. .546
FT /label= Extracellular_domain
FT Peptide 1. .20
FT /label= Signal_peptide
FT Protein 21. .852
FT /label= Mature_GPCR50289
FT Domain 61. .470
FT /label= ANF_domain
FT /note= "Natriuretic peptide (ANF) ligand binding domain"
XX
PN WO200164882-A2.
XX
PD 07-SEP-2001.
XX
PF 28-FEB-2001; 2001WO-US006543.
XX
PR 29-FEB-2000; 2000US-0186059P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Glucksmann MA, Galvin KM, Silos-Santiago I;
XX
DR WPI; 2001-589866/66.
DR N-PSDB; AAS14574.
XX
PT Novel G protein coupled receptors and nucleic acids encoding them, for
PT identifying agents for the treatment of cardiac disorders.
XX
PS Claim 9; Fig 15; 209pp; English.
XX
CC The invention relates to novel human G protein-coupled receptors (GPCR)
CC named 1983, 52881, 2398, 45449, 50289 and 52872. The novel GPCRs and
CC nucleic acids encoding them are useful for identifying agents for the
CC treatment of cardiovascular disorders, angiogenesis-related disorders,
CC neural disorders, pain response disorders and inflammatory disorders e.g.
CC atherosclerosis, angina pectoris and myocardial infarction, ischaemic
CC heart disease, sudden cardiac death, hypertensive heart disease,
CC diabetes, prostate cancer-related pain, diabetes and obesity. The present
CC sequence represents GPCR 50289
XX
SQ Sequence 852 AA;

Query Match 99.7%; Score 4512; DB 4; Length 852;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 851; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360

Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420

Qy        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540

Qy        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600

Qy        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720

Qy        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGRYNRARGLTFAMLAYFITWVSFVP 780

Qy        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840

Qy        841 NDGNTGNQGKHE 852
              ||||||||||||
Db        841 NDGNTGNQGKHE 852
```

RESULT 7
AAO21501
ID AAO21501 standard; protein; 852 AA.
XX
AC AAO21501;

```
XX
DT   15-AUG-2002  (first entry)
XX
DE   Human G-protein coupled receptor (GPCR).
XX
KW   G-protein coupled receptor; GPCR; gene chip; human; immune response;
KW   chromosomal position; transgenic animal; gene therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Region          20. .25
FT                   /note= "Myristoylation site"
FT   Region          69. .74
FT                   /note= "Myristoylation site"
FT   Modified-site   85. .88
FT                   /note= "Asn is N-glycosylated"
FT   Region          92. .97
FT                   /note= "Myristoylation site"
FT   Region          102. .105
FT                   /note= "Casein kinase II phosphorylation site"
FT   Modified-site   130. .133
FT                   /note= "Asn is N-glycosylated"
FT   Domain          137. .167
FT                   /label= Helix_1
FT                   /note= "Membrane spanning structure and domain"
FT   Region          153. .155
FT                   /note= "Protein kinase C phosphorylation site"
FT   Region          175. .178
FT                   /note= "Casein kinase II phosphorylation site"
FT   Region          175. .177
FT                   /note= "Protein kinase C phosphorylation site"
FT   Region          189. .191
FT                   /note= "Protein kinase C phosphorylation site"
FT   Region          214. .217
FT                   /note= "Casein kinase II phosphorylation site"
FT   Region          234. .239
FT                   /note= "Myristoylation site"
FT   Domain          256. .286
FT                   /label= Helix_3
FT                   /note= "Membrane spanning structure and domain"
FT   Modified-site   264. .267
FT                   /note= "Asn is N-glycosylated"
FT   Modified-site   285. .288
FT                   /note= "Asn is N-glycosylated"
FT   Region          289. .291
FT                   /note= "Protein kinase C phosphorylation site"
FT   Region          293. .295
FT                   /note= "Protein kinase C phosphorylation site"
FT   Domain          306. .326
FT                   /label= Helix_4
FT                   /note= "Membrane spanning structure and domain"
FT   Region          319. .324
FT                   /note= "Myristoylation site"
FT   Modified-site   380. .383
FT                   /note= "Asn is N-glycosylated"
FT   Modified-site   411. .414
FT                   /note= "Asn is N-glycosylated"
FT   Modified-site   432. .435
FT                   /note= "Asn is N-glycosylated"
FT   Modified-site   475. .478
FT                   /note= "Asn is N-glycosylated"
FT   Region          476. .481
FT                   /note= "Myristoylation site"
FT   Region          477. .479
FT                   /note= "Protein kinase C phosphorylation site"
FT   Region          480. .482
FT                   /note= "Protein kinase C phosphorylation site"
FT   Region          517. .541
FT                   /note= "G-protein coupled receptor family 3 signature 2"
FT   Region          528. .530
FT                   /note= "Protein kinase C phosphorylation site"
FT   Region          551. .553
FT                   /note= "Protein kinase C phosphorylation site"
FT   Region          556. .559
FT                   /note= "cGMP-dependent protein kinase phosphorylation
FT                   site"
FT   Domain          567. .587
FT                   /label= Helix_6
FT                   /note= "Membrane spanning structure and domain"
FT   Region          581. .586
FT                   /note= "Myristoylation site"
FT   Region          603. .608
FT                   /note= "Myristoylation site"
FT   Domain          606. .626
FT                   /label= Helix_7
FT                   /note= "Membrane spanning structure and domain"
FT   Domain          637. .657
FT                   /label= Helix_8
FT                   /note= "Membrane spanning structure and domain"
FT   Region          646. .651
FT                   /note= "Myristoylation site"
FT   Region          667. .670
FT                   /note= "Casein kinase II phosphorylation site"
FT   Domain          680. .700
FT                   /label= Helix_9
FT                   /note= "Membrane spanning structure and domain"
FT   Domain          731. .751
FT                   /label= Helix_10
FT                   /note= "Membrane spanning structure and domain"
FT   Region          731. .736
FT                   /note= "Myristoylation site"
FT   Modified-site   737. .740
FT                   /note= "Asn is N-glycosylated"
FT   Domain          763. .783
FT                   /label= Helix_11
FT                   /note= "Membrane spanning structure and domain"
FT   Region          763. .768
FT                   /note= "Myristoylation site"
FT   Domain          792. .812
FT                   /label= Helix_12
FT                   /note= "Membrane spanning structure and domain"
FT   Region          804. .809
FT                   /note= "Myristoylation site"
FT   Region          831. .836
FT                   /note= "Myristoylation site"
FT   Region          839. .844
FT                   /note= "Myristoylation site"
FT   Region          843. .848
FT                   /note= "Myristoylation site"
XX
PN   WO200230981-A1.
XX
PD   18-APR-2002.
XX
PF   13-MAR-2001; 2001WO-US007832.
XX
PR   10-OCT-2000; 2000US-00684393.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Wei M,  Zhong W,  Ketchum KA,  Difrancesco V,  Beasley EM;
XX
DR   WPI; 2002-444173/47.
DR   N-PSDB; AAL38462, AAL38463.
XX
PT   Novel G protein coupled receptor, useful for raising antibodies, to
PT   elicit immune response, and as a reagent in assays designed to
PT   quantitatively determine levels of protein in biological samples.
XX
PS   Claim 1; Fig 2; 82pp; English.
XX
CC   The invention relates to an isolated G-protein coupled receptor (GPCR)
CC   polypeptide, comprising an 852 residue amino acid sequence, given in the
CC   specification, an allelic variant or ortholog of the protein, or a
CC   fragment comprising at least 10 contiguous amino acids of the protein.
```

```
CC   GPCR is useful for identifying a modulator of GPCR and an agent that
CC   binds to GPCR. GPCR and a gene chip comprising GPCR are useful as models
CC   for the development of human therapeutic agents. GPCR is useful for
CC   raising antibodies, to elicit immune response, as a reagent in assays
CC   designed to quantitatively determine levels of protein in biological
CC   samples, and as markers for tissues in which the corresponding protein is
CC   preferentially expressed. A gene chip containing GPCR is also useful as a
CC   probe for determining the chromosomal positions of nucleic acid molecules
CC   by means of in situ hybridisation, in making vectors containing the gene
CC   regulatory regions of a gene chip containing GPCR, for designing
CC   ribozymes, in making vectors that express GPCR, and for constructing host
CC   cells and transgenic animals expressing nucleic acid molecules and
CC   peptides. A host cell containing GPCR is useful for conducting cell-based
CC   assays involving GPCR protein or its fragments, and for identifying
CC   protein mutants in which these function is affected. The polynucleotide
CC   encoding GPCR can be used to treat disorders by gene therapy. This
CC   sequence represents the human G-protein coupled receptor of the invention
XX
SQ   Sequence 852 AA;

  Query Match             99.7%;  Score 4512;  DB 5;  Length 852;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  851;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360

Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420

Qy        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540

Qy        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600

Qy        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720

Qy        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGRYNRARGLTFAMLAYFITWVSFVP 780

Qy        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840

Qy        841 NDGNTGNQGKHE 852
              ||||||||||||
Db        841 NDGNTGNQGKHE 852


RESULT 8
AAU78587
ID   AAU78587 standard; protein; 852 AA.
XX
AC   AAU78587;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Human AXOR79 protein.
XX
KW   AXOR79; human; seven transmembrane; G-protein coupled; receptor;
KW   bacterial; viral; fungal; infection; protozoal; manic depression;
KW   human immunodeficiency virus; HIV; cancer; diabetes; anorexia; bulimia;
KW   Parkinson's disease; heart failure; asthma; allergy; osteoporosis;
KW   neurological disorder; schizophrenia; Huntington's disease.
XX
OS   Homo sapiens.
XX
PN   GB2364058-A.
XX
PD   16-JAN-2002.
XX
PF   03-MAY-2001; 2001GB-00010887.
XX
PR   05-MAY-2000; 2000US-00566161.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (SMIK ) SMITHKLINE BEECHAM PLC.
XX
PI   Ellis CE,  Gattu M;
XX
DR   WPI; 2002-306799/35.
DR   N-PSDB; ABK47345.
XX
PT   New G protein-coupled receptor AXOR79, useful for diagnosis and treatment
PT   of e.g. infections and cancer.
XX
PS   Claim 1; Page 28; 36pp; English.
XX
CC   This invention relates to the nucleotide and protein sequences of a novel
CC   isolated G-protein coupled 7 transmembrane receptor AXOR79. The DNA and
CC   protein sequences of the invention are useful for treating a wide range
CC   of diseases such as infections (bacterial, viral, fungal or protozoal,
CC   especially human immunodeficiency virus (HIV)-1 or -2), pain, cancer,
CC   diabetes, anorexia, bulimia, Parkinson's disease, heart failure, asthma,
CC   allergy, osteoporosis, neurological disorders (e.g. schizophrenia, manic
CC   depression, Huntington's disease) and many others). The sequences are
CC   also useful to diagnose, and predict susceptibility to, diseases
CC   including detecting mutations or measuring mRNA or protein levels
CC   associated with inappropriate activity or levels of AXOR79. The sequences
CC   may also be used to screen for agonists or antagonists of AXOR79 and for
CC   potential therapeutic agents. The nucleotide sequence can also be used to
CC   produce recombinant AXOR79 protein which may be used as a vaccine or as
CC   source of primers and probes for isolation of related nucleic acids, for
CC   chromosomal localisation and in tissue expression studies. The protein
CC   sequence can be used to raise specific antibodies which are useful as
CC   diagnostic reagents, for isolation/identification of expression clones,
CC   for affinity purification of the protein and also for treating the
CC   specified diseases. The present sequence represents the human AXOR79 G-
CC   protein coupled 7 transmembrane receptor protein of the invention
XX
```

```
SQ   Sequence 852 AA;

  Query Match           99.7%;  Score 4512;  DB 5;  Length 852;
  Best Local Similarity 99.9%;  Pred. No. 0;
  Matches  851;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360

Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420

Qy        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540

Qy        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600

Qy        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720

Qy        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGRYNRARGLTFAMLAYFITWVSFVP 780

Qy        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840

Qy        841 NDGNTGNQGKHE 852
              ||||||||||||
Db        841 NDGNTGNQGKHE 852


RESULT 9
ABP70451
ID   ABP70451 standard; protein; 852 AA.
XX
AC   ABP70451;
XX
DT   22-APR-2003  (first entry)
XX
DE   Amino acid sequence of human T1R3.
XX
KW   T1 receptor; T1R3; G protein-coupled receptor; GPCR; T1R2; T1R3;
KW   sweet taste stimuli; T1R2; umami taste stimuli; taste stimuli;
KW   taste perception; taste sensation.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 757
FT                   /note= "Arg encoded by TGC"
XX
PN   WO2003001876-A2.
XX
PD   09-JAN-2003.
XX
PF   26-JUN-2002; 2002WO-US019970.
XX
PR   26-JUN-2001; 2001US-0300434P.
PR   03-JUL-2001; 2001US-00897427.
PR   13-JUL-2001; 2001US-0304749P.
PR   08-AUG-2001; 2001US-0310493P.
PR   21-NOV-2001; 2001US-0331771P.
PR   14-DEC-2001; 2001US-0339472P.
PR   03-JAN-2002; 2002US-00035045.
PR   15-APR-2002; 2002US-0372090P.
PR   22-APR-2002; 2002US-0374143P.
XX
PA   (SENO-) SENOMYX INC.
XX
PI   Zoller MT,  Li X,  Staszewski L,  O'connell S,  Zozulya S,  Adler JE;
PI   Xu H,  Echeverri F;
XX
DR   WPI; 2003-210181/20.
DR   N-PSDB; ABZ68280.
XX
PT   New receptor that comprises T1R1 and/or T1R3 or T1R2 and/or T1R3
PT   polypeptide and that specifically binds to and/or is activated by umami
PT   or sweet taste stimuli, useful for identifying compounds that modulate
PT   taste perception.
XX
PS   Example 1; Page 80-81; 135pp; English.
XX
CC   The present sequence represents a human T1 receptor designated T1R3. The
CC   receptor is a G protein-coupled receptor (GPCR). T1R2 and T1R3 act in
CC   combination to recognise sweet taste stimuli, while T1R2 and T1R3 act in
CC   combination to recognise umami taste stimuli. The specification describes
CC   different combinations of T1Rs, that when co-expressed, produce
CC   functional taste receptors that respond to taste stimuli. The hetero-
CC   oligomeric taste receptors of the invention are useful for identifying
CC   compounds that modulate taste perception. They are also useful for
CC   modifying taste sensation in an animal
XX
SQ   Sequence 852 AA;

  Query Match           99.7%;  Score 4512;  DB 6;  Length 852;
  Best Local Similarity 99.9%;  Pred. No. 0;
  Matches  851;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
```

```
Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360

Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420

Qy        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540

Qy        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600

Qy        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720

Qy        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGRYNRARGLTFAMLAYFITWVSFVP 780

Qy        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840

Qy        841 NDGNTGNQGKHE 852
              ||||||||||||
Db        841 NDGNTGNQGKHE 852
```

RESULT 10
ABU62106
ID ABU62106 standard; protein; 852 AA.
XX
AC ABU62106;
XX
DT 19-AUG-2003 (first entry)
XX
DE Human calcium sensing G-protein coupled receptor.
XX
KW Human; receptor; GPCR; G-protein coupled receptor; gene therapy;
KW Parkinson's disease; Huntington's disease; Tourette's syndrome.
XX
OS Homo sapiens.
XX
PN US2003036089-A1.
XX
PD 20-FEB-2003.
XX
PF 02-OCT-2002; 2002US-00261482.
XX
PR 20-DEC-1999; 99US-0172600P.
PR 10-OCT-2000; 2000US-00684393.
XX
PA (APPL-) APPLERA CORP.
XX
PI Wei M, Zhong W, Ketchum KA, Di Francesco V, Beasley EM;
XX
DR WPI; 2003-492117/46.
DR N-PSDB; ACA62666, ACA62667.
XX
PT New isolated human G-protein coupled receptor proteins, useful for
PT treating e.g. Parkinson's disease, Huntington's disease or Tourette's
PT syndrome, or for developing human therapeutics and diagnostic
PT compositions.
XX
PS Claim 1; Fig 2A; 46pp; English.
XX
CC The invention relates to an isolated human G-protein coupled receptor
CC (GPCR) peptide. The human GPCR peptides and nucleic acid molecules are
CC useful in the development of human therapeutics and diagnostic
CC compositions. The peptides are useful for treating disorders (e.g.
CC Parkinson's disease, Huntington's disease or Tourette's syndrome)
CC characterised by an absence of, inappropriate, or unwanted expression of
CC the GPCR protein, or disorders mediated by proteases. These molecules are
CC particularly useful as models for developing human therapeutic targets,
CC identifying therapeutic proteins, or serving as targets for the
CC development of human therapeutic agents that modulate GPCR activity in
CC cells and tissues that express the GPCR. The peptides are also useful for
CC raising antibodies or eliciting an immune response, as a reagent
CC (including the labelled reagent) in assays designed to quantitatively
CC determine levels of the protein (or its binding partner or ligand) in
CC biological fluids, or as markers for tissues in which the corresponding
CC protein is preferentially expressed. The agents identified are useful for
CC treating a subject with a disorder mediated by human proteases. The
CC present sequence represents the amino acid sequence of human calcium
CC sensing G-protein coupled receptor
XX
SQ Sequence 852 AA;

Query Match 99.7%; Score 4512; DB 6; Length 852;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 851; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360

Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420

Qy        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540

Qy        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600

Qy        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720

Qy        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
                |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGRYNRARGLTFAMLAYFITWVSFVP 780

Qy        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840

Qy        841 NDGNTGNQGKHE 852
                ||||||||||||
Db        841 NDGNTGNQGKHE 852
```

```
RESULT 11
ADG19752
ID   ADG19752 standard; protein; 852 AA.
XX
AC   ADG19752;
XX
DT   26-FEB-2004  (first entry)
XX
DE   Human G protein coupled receptor (GPCR) 50289.
XX
KW   cytostatic; nephrotropic; antiinflammatory; cardiant; respiratory;
KW   gastrointestinal; neuroprotective; angiogenesis stimulator; gene therapy;
KW   18636; 2466; 43238; 1983; 52881; 2398; 45449; 50289; 52872; 26908;
KW   cellular proliferative disorder; brain disorder; renal disorder;
KW   kidney disorder; inflammatory disorder; cardiovascular disorder;
KW   liver disorder; intestinal disorder; respiratory disorder;
KW   angiogenic disorder; human; G protein coupled receptor; GPCR.
XX
OS   Homo sapiens.
XX
PN   US2003215860-A1.
XX
PD   20-NOV-2003.
XX
PF   03-APR-2003; 2003US-00407079.
XX
PR   29-FEB-2000; 2000US-0186059P.
PR   24-MAR-2000; 2000US-0191845P.
PR   22-MAY-2000; 2000US-0206019P.
PR   17-NOV-2000; 2000US-00715790.
PR   28-FEB-2001; 2001US-00796338.
PR   22-MAY-2001; 2001US-00863200.
PR   22-AUG-2001; 2001US-0314041P.
PR   22-AUG-2001; 2001US-0314185P.
PR   21-AUG-2002; 2002US-00225094.
PR   22-AUG-2002; 2002US-00226102.
PR   15-OCT-2002; 2002US-00272417.
PR   29-OCT-2002; 2002US-00282837.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Glucksmann MA,  Silos-Santiago I,  Carroll JM,  Galvin KM;
XX
DR   WPI; 2004-010777/01.
DR   N-PSDB; ADG19751, ADG19753.
XX
PT   New nucleic acid or polypeptide 18636, 2466, 43238, 1983, 52881, 2398,
PT   45449, 50289, 52872 or 26908, useful in preparing a composition for
PT   treating e.g., cellular proliferative, brain, kidney, inflammatory or
PT   angiogenic disorder.
XX
PS   Claim 4; SEQ ID NO 90; 163pp; English.
XX
CC   The invention describes a new isolated 18636, 2466, 43238, 1983, 52881,
CC   2398, 45449, 50289, 52872 or 26908 nucleic acid molecule comprising: a
CC   sequence comprising 939-3489 bp or its fragment comprising at least 15
CC   nucleotides; a sequence encoding a polypeptide comprising a sequence
CC   having 223-852 amino acids, or its allelic variant or fragment comprising
CC   at least 15 contiguous amino acids; or a sequence that hybridises with
CC   (1) under stringent conditions. The nucleic acid or polypeptide is useful
CC   in preparing a composition for treating a disorder characterised by
CC   aberrant 18636, 2466, 43238, 1983, 52881, 2398, 45449, 50289, 52872 or
CC   26908 activity e.g., cellular proliferative, brain, renal, kidney,
CC   inflammatory, cardiovascular, liver, intestinal, respiratory or
CC   angiogenic disorder. This is the amino acid sequence of novel human
CC   protein 50289, a G protein coupled receptor (GPCR).
XX
SQ   Sequence 852 AA;

  Query Match             99.7%;  Score 4512;  DB 8;  Length 852;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  851;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360

Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420

Qy        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540

Qy        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600

Qy        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720

Qy        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
                |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGRYNRARGLTFAMLAYFITWVSFVP 780

Qy        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840

Qy        841 NDGNTGNQGKHE 852
                ||||||||||||
Db        841 NDGNTGNQGKHE 852


RESULT 12
ADG73581
ID   ADG73581 standard; protein; 852 AA.
XX
AC   ADG73581;
XX
DT   11-MAR-2004  (first entry)
XX
DE   Human T1R3.
XX
KW   umami receptor; sweet taste receptor; taste perception; food; beverage;
KW   umami; sweet taste; human.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 757
FT                   /note= "Encoded by TGC"
XX
PN   US2003232407-A1.
XX
PD   18-DEC-2003.
XX
PF   26-JUN-2002; 2002US-00179373.
XX
PR   26-JUN-2001; 2001US-0300434P.
PR   13-JUL-2001; 2001US-0304749P.
PR   08-AUG-2001; 2001US-0310493P.
PR   21-NOV-2001; 2001US-0331771P.
PR   14-DEC-2001; 2001US-0339472P.
PR   15-APR-2002; 2002US-0372090P.
PR   22-APR-2002; 2002US-0374143P.
PR   23-APR-2002; 2002US-0374522P.
XX
PA   (SENO-) SENOMYX INC.
XX
PI   Zoller M,  Li X,  Staszewski L,  O'connell S,  Zozulya S,  Adler JE;
PI   Xu H,  Echeverri F;
XX
DR   WPI; 2004-052170/05.
DR   N-PSDB; ADG73583.
XX
PT   Novel receptors T1R1 and T1R3 that respond to umami taste stimuli and
PT   receptors T1R2 and T1R3 that respond to sweet taste stimuli, useful for
PT   identifying modulators which are used for modifying taste sensation in
PT   animals.
XX
PS   Example 1; SEQ ID NO 7; 55pp; English.
XX
CC   The invention relates to a receptor comprised of at least one T1R1
CC   polypeptide and/or at least one T1R3 polypeptide or a receptor comprised
CC   of at least one T1R2 polypeptide and/or at least one T1R3 polypeptide
CC   where the receptors specifically bind to and/or are activated by umami or
CC   sweet taste stimuli. The receptor is useful for identifying compounds
CC   that modulate taste perception, where the method involves identifying
CC   compounds that bind to, activate, inhibit, enhance and/or modulate one or
CC   more of the receptors. The receptor is useful for quantifying the taste
CC   of individual compounds or food or beverage compositions. The method is
CC   used to screen a compound library for compounds that enhance or modulate
CC   the activity of L-glutamate to activate the T1R1/T1R3 umami taste
CC   receptor, or to screen a compound library for compounds that agonise or
CC   antagonise the T1R1/T1R3 sweet taste receptor. The method screens for a
CC   compound that competes with IMP, GMP or their analogues for binding to
CC   the T1R1/T1R3 umami taste receptor. The method is used to screen a
CC   compound library for compounds that mimic the activity of IMP, GMP or
CC   their analogues that enhance the activity of a T1R1/T1R3 agonist or is
CC   used to screen a compound library for compounds that enhance or modulate
CC   the activity of a sweetener to activate the T1R2/T1R3 sweet taste
CC   receptor. The method is useful for modifying umami taste or sweet taste
CC   sensation in an animal such as human, dog, cat, fish, cow, sheep or pig.
CC   The compound is formulated in a food, beverage, or oral pharmaceutical
CC   composition. The present sequence represents the amino acid sequence of
CC   human T1R3.
XX
SQ   Sequence 852 AA;

  Query Match             99.7%;  Score 4512;  DB 8;  Length 852;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  851;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360

Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420

Qy        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540

Qy        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600

Qy        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720

Qy        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
```

```
Db        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGRYNRARGLTFAMLAYFITWVSFVP 780

Qy        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840

Qy        841 NDGNTGNQGKHE 852
              ||||||||||||
Db        841 NDGNTGNQGKHE 852
```

RESULT 13
ADR29260
ID  ADR29260 standard; protein; 852 AA.
XX
AC  ADR29260;
XX
DT  04-NOV-2004  (first entry)
XX
DE  Taste receptor modulation-related human T1R3 protein sequence SeqID199.
XX
KW  T1R; T2R; taste receptor; G protein; modulatory compound; MAPK activity;
KW  cAMP accumulation; adenylyl cyclase; bitter; taste; umami; sweet; T1R3;
KW  human.
XX
OS  Homo sapiens.
XX
PN  WO2004069191-A2.
XX
PD  19-AUG-2004.
XX
PF  03-FEB-2004; 2004WO-US002987.
XX
PR  03-FEB-2003; 2003US-0444172P.
PR  26-MAR-2003; 2003US-0457318P.
XX
PA  (SENO-) SENOMYX INC.
XX
PI  Servant G,  Ozeck M,  Brust P,  Xu H;
XX
DR  WPI; 2004-604341/58.
DR  N-PSDB; ADR29262.
XX
PT  Identifying a compound that modulates the activity of a T1R or T2R taste
PT  receptor by assaying the effect of the putative modulatory compound on
PT  MAPK activation, cAMP accumulation or adenylyl cyclase activity in the
PT  eukaryotic cell.
XX
PS  Disclosure; SEQ ID NO 199; 248pp; English.
XX
CC  This invention is related to a novel method of identifying a compound
CC  that modulates the activity of a T1R or T2R taste receptor. The method
CC  comprises providing a eukaryotic cell that expresses a functional T1R or
CC  T2R taste receptor and a G protein that couples to it, contacting the
CC  eukaryotic cell with a compound and identifying whether the compound
CC  modulates the activity of T1R or T2R expressed by the eukaryotic cell
CC  based on its effect on MAPK activity, cAMP accumulation or adenylyl
CC  cyclase activity. The method is used to identify a compound that blocks
CC  bitter taste associated with a particular T2R activator or that blocks or
CC  enhances umami taste elicited by a compound that activates the T1R1/T1R3
CC  (umami) taste receptor or sweet taste elicited by a compound that
CC  activates the T1R2/T1R3 (sweet) taste receptor. The method is useful in
CC  identifying a compound that modulates the activity of a T1R or T2R taste
CC  receptor. The present sequence is that of a human T1R taste receptor
CC  protein which is related to the method of the invention.
XX
SQ  Sequence 852 AA;

```
  Query Match             99.7%;  Score 4512;  DB 8;  Length 852;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  851;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360

Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420

Qy        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540

Qy        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600

Qy        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720

Qy        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGRYNRARGLTFAMLAYFITWVSFVP 780

Qy        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840

Qy        841 NDGNTGNQGKHE 852
              ||||||||||||
Db        841 NDGNTGNQGKHE 852
```

RESULT 14
ADW26968
ID  ADW26968 standard; protein; 852 AA.
XX
AC  ADW26968;
XX
DT  07-APR-2005  (first entry)
XX
DE  Human taste receptor, T1R3.
XX
KW  Taste receptor; Tas1r3; T1R3; diabetes mellitus; obesity;
KW  nutritional disorder; anorectic; antidiabetic; receptor.
XX

```
OS   Homo sapiens.
XX
PN   WO2005005480-A2.
XX
PD   20-JAN-2005.
XX
PF   13-MAY-2004; 2004WO-US015136.
XX
PR   27-JUN-2003; 2003US-0482992P.
PR   19-MAR-2004; 2004US-0554751P.
XX
PA   (MONE-) MONELL CHEM SENSES CENT.
PA   (LIXX/) LI X.
PA   (LIWW/) LI W.
PA   (REED/) REED D R.
PA   (BACH/) BACHMANOV A A.
PA   (BRAN/) BRAND J G.
XX
PI   Li X,  Li W,  Reed DR,  Bachmanov AA,  Brand JG;
XX
DR   WPI; 2005-092067/10.
DR   N-PSDB; ADW26967.
XX
PT   Novel isolated and purified polynucleotide encoding feline T1R1, T1R2 or
PT   T1R3 receptor, useful for identifying agonist or antagonist of feline T1R
PT   receptor and for modifying taste perception of mammal.
XX
PS   Example; SEQ ID NO 12; 183pp; English.
XX
CC   The invention relates to the isolation of polynucleotide sequences
CC   (Tas1r1, Tas1r2, Tas1r3) encoding taste receptors of the T1R family, and
CC   the polypeptide sequences they encode (T1R1, T1R2, T1R3). Also described
CC   are (i) an expression vector comprising a Tas1r polynucleotide sequence
CC   operably linked to a promoter, (ii) a host cell comprising the expression
CC   vector, (iii) a cell culture comprising the host cell, (iv) a kit for the
CC   detection of a polynucleotide encoding a feline T1R polypeptide, (v) a
CC   feline T1R polypeptide produced using the host cell above, (vi) a method
CC   of identifying a feline T1R polypeptide variant that confers modified
CC   taste perception, (vii) a transgenic animal comprising a heterologous
CC   feline T1R allele, (viii) an isolated antibody that specifically
CC   immunoreacts with at least one epitope of a feline T1R polypeptide or
CC   dimmer, (ix) a method of identifying compounds that interact with a
CC   feline T1R polypeptide dimer, and (x) a method of identifying an agonist
CC   or antagonist of a feline T1R polypeptide dimer. The polynucleotide and
CC   polypeptide sequences of the invention are useful for identifying an
CC   agonist or antagonist of a feline T1R receptor. The sequences and methods
CC   are also useful for the treatment of disorders such as diabetes and
CC   obesity in mammals such as avian, feline, canine, bovine, ovine, porcine,
CC   equine, rodent, simian or humans. This sequence represents a human taste
CC   receptor.
XX
SQ   Sequence 852 AA;

  Query Match             99.7%;  Score 4512;  DB 9;  Length 852;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  851;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360

Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420

Qy        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540

Qy        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600

Qy        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720

Qy        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGRYNRARGLTFAMLAYFITWVSFVP 780

Qy        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840

Qy        841 NDGNTGNQGKHE 852
              ||||||||||||
Db        841 NDGNTGNQGKHE 852


RESULT 15
ADY27327
ID   ADY27327 standard; protein; 852 AA.
XX
AC   ADY27327;
XX
DT   05-MAY-2005  (first entry)
XX
DE   Synthetic T1R construct SEQ ID NO 7.
XX
KW   T1R2 receptor; T1R3 receptor; T1R1 receptor; savory; sweet; taste; umami;
KW   food; beverage; medicine; protein engineering.
XX
OS   unidentified.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 757
FT                   /note= "Encoded by TGC"
XX
PN   WO2005015158-A2.
XX
PD   17-FEB-2005.
XX
PF   06-AUG-2004; 2004WO-US025459.
XX
PR   06-AUG-2003; 2003US-0494071P.
PR   09-MAR-2004; 2004US-0552064P.
```

XX
PA (SENO-) SENOMYX INC.
XX
PI Li X, Staszewski L, Xu H;
XX
DR WPI; 2005-163285/17.
DR N-PSDB; ADY27329.
XX
PT New non-naturally occurring compound that specifically binds to a
PT T1R2/T1R3 receptor, useful for modulating the sweet taste of a desired
PT food, beverage or medicinal.
XX
PS Disclosure; SEQ ID NO 7; 244pp; English.
XX
CC This invention describes a novel non-naturally occurring compound that
CC specifically binds to a T1R2/T1R3 receptor composed of hT1R2/hT1R3 but
CC not rT1R2/rT1R3; T1R2/T1R3 receptor composed of hT1R2/rT1R3 but not
CC rT1R2/hT1R3; N-terminal extracellular domain of T1R2 of the hT1R2/hT1R3
CC receptor or of T1R1 of the T1R1/hT1R3 receptor; T1R2/T1R3 receptor
CC composed of rT1R2/hT1R3 but not hT1R2/rT1R3; hT1R2/hT1R3 and rT1R2/r3-h3
CC but not to rT1R2/rT1R3 or to hT1R2/h3-r3; hT1R2/hT1R3 and r2-h2/rT1R3 but
CC not to rT1R2/rT1R3 or to h2-r2/hT1R3; Venus Flytrap Domain (VFD) of T1R2
CC of the hT1R2/hT1R3 and hT1R2/rT1R3 receptor; human N-terminal VFD
CC residues 144 to 302 of the T1R2 subunit of the T1R2/T1R3 receptor; N-
CC terminal VFD of the T1R2 subunit of the T1R2/T1R3 receptor; cysteine-rich
CC region of T1R2 of the hT1R2/hT1R3 receptor or of T1R1 of the T1R1/hT1R3
CC receptor; Transmembrane Domain (TM) of T1R2 of the hT1R2/hT1R3 receptor;
CC human N-terminal extracellular domain of the T1R3 subunit of the
CC T1R2/T1R3; VFD of T1R3 of the hT1R2/hT1R3 receptor; TM of T1R3 of the
CC hT1R2/hT1R3 receptor; extracellular loop 2 and extracellular loop 3 of
CC the human transmembrane domain of the T1R3 subunit of T1R2/T1R3; T1R1VFD
CC of the T1R1/T1R3 savory receptor; T1R1 TM domain of the T1R1/T1R3 savory
CC receptor or T1R3 TM domain of the T1R1/T1R3 savory receptor; N-terminal
CC extracellular domain of T1R3 of the T1R1/hT1R3 receptor; T1R3 VFD of the
CC T1R1/T1R3 savory receptor; cysteine-rich region of T1R3 of the T1R1/hT1R3
CC receptor; TM domain of T1R1 of a truncated savory receptor composed of
CC the h1TM/h3TM or TM domain of T1R3 of a truncated sweet receptor composed
CC of h1TM/h3TM; or TM domain of T1R1 or T1R3 of a chimeric receptor
CC composed of mGluR-h1/mGluR-h3. The invention also describes a method for
CC identifying compounds that modulate taste perception, a method for
CC identifying a compound that modulates sweet taste perception or umami
CC taste perception, a method for modulating or inhibiting or increasing the
CC savory taste or sweet taste of a comestible or medicinal product and a
CC method of enhancing sweet taste or umami taste perception. The compound
CC of the invention demonstrates compound-dependent increase in fluorescence
CC with an activity compared to the maximal activity for fructose or
CC glutamate of at least 25% in a fluorescence-based assay using a FLIPR
CC (Molecular Devices) instrument. Enhancing sweet or umami taste perception
CC comprises contacting a sweet or umami receptor with cyclamate and NHDC,
CC and their derivatives or with lactisole derivatives. The compound is
CC useful for modulating the sweet taste of a desired food, beverage or
CC medicine. The chimeric receptors are useful in assays to identify
CC compounds that respectively respond to umami taste stimuli and sweet
CC taste stimuli. This sequence represents a synthetic T1R construct
CC represented in the disclosure of the invention.
XX
SQ Sequence 852 AA;

Query Match 99.7%; Score 4512; DB 9; Length 852;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 851; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360

Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420

Qy        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540

Qy        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600

Qy        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720

Qy        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGRYNRARGLTFAMLAYFITWVSFVP 780

Qy        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840

Qy        841 NDGNTGNQGKHE 852
              ||||||||||||
Db        841 NDGNTGNQGKHE 852
```

Search completed: December 16, 2005, 18:06:37
Job time : 198 secs

THIS PAGE BLANK (USPTO)

;  FILING DATE: 21 October, 1994
;  APPLICATION NUMBER: U.S. 08/292,827
;  FILING DATE: 23 August, 1994
;  APPLICATION NUMBER: U.S. 08/141,248
;  FILING DATE: 22 October, 1993
;  APPLICATION NUMBER: U.S. 08/009,389
;  FILING DATE: 23 February, 1993
;  APPLICATION NUMBER: U.S. 08/017,127
;  FILING DATE: 12 February, 1993
;  APPLICATION NUMBER: U.S. 07/934,161
;  FILING DATE: 21 August, 1992
;  APPLICATION NUMBER: U.S. 07/834,044
;  FILING DATE: 11 February, 1992
;  APPLICATION NUMBER: U.S. 07/749,451
;  FILING DATE: 23 August, 1991
;  ATTORNEY/AGENT INFORMATION:
;  NAME: Heber, Sheldon O.
;  REGISTRATION NUMBER: 38,179
;  REFERENCE/DOCKET NUMBER: 213/006
;  TELECOMMUNICATION INFORMATION:
;  TELEPHONE: (213) 489-1600
;  TELEFAX: (213) 955-0440
;  TELEX: 67-3510
;  INFORMATION FOR SEQ ID NO: 5:
;  SEQUENCE CHARACTERISTICS:
;  LENGTH: 1085 amino acids
;  TYPE: amino acid
;  TOPOLOGY: linear
;  MOLECULE TYPE: protein
US-08-484-565-5

Query Match 23.4%; Score 1060; DB 1; Length 1085;
Best Local Similarity 30.0%; Pred. No. 1.7e-92;
Matches 269; Conservative 155; Mismatches 362; Indels 112; Gaps 26;

```
Qy          6 VLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPL--GEA-EEAGLRSRTRPSSPVC 62
              :|  | |      |         |: : ||| :||||||:  | | ::  |:|  || |  |
Db         10 LLAFSTWCTSAYGP------DQRAQKKGDIILGGLFPIHFGVAVKDQDLKS--RPESVEC 61

Qy         63 TRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLA--KA 120
              |::  |  |  ||  |:||||:   ||| : ||| :||||:    |:: :| |:|  |
Db         62 IRYNFRGFRWLQAMIFAIEEINSSPALLPNMTLGYRIFDTCNTVSKALEATLSFVAQNKI 121

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              | ::  :|| ::: |  :||:|   | ::         | :||||| :|  ||| :  |
Db        122 DSLNLDEFCNCSEHIPSTIAVVGATGSGISTAVANLLGLFYIPQVSYASSSRLLSNKNQF 181

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              || ||:|:|  | || |:::: | ||||  : :||:||| |:  |   |  | |||
Db        182 KSFLRTIPNDEHQATAMADIIEYFRWNWVGTIAADDDYGRPGIEKFREEAEERDICIDFS 241

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              |:   :  |  |:| |:  :  |: :|:::|:|      |   :  :: ::|:||
Db        242 ELI----SQYSDEEKIQQVVEVIQNSTAKVIVVFSSGPDLEPLIKEIVRRNITGRIWLAS 297

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYV-KTHLALATDPAFCSAL----- 354
              ||| :| |:        :|  :||  :  |:  | ::: | |   :    |
Db        298 EAWASSSLIAMPEYFHVVGGTIGFGLKAGQIPGFREFLQKVHPRKSVHNGFAKEFWEETF 357

Qy        355 ------------------GEREQGL---------------EEDVVGQRCPQCDCITLQN 380
                             |  | |                ||::     |  |   |:
Db        358 NCHLQEGAKGPLPVDTFLRGHEEGGARLSNSPTAFRPLCTGEENISSVETPYMDYTHLR- 416

Qy        381 VSAGLNHHQTFSVYAAVYSVAQALHNTLQC--------NASGCPAQDPVKPWQLLENMYN 432
                      :::|| ||||:| || :   |         | | |      |: ||:|::: :
Db        417 --------ISYNVYLAVYSIAHALQDIYTCIPGRGLFTNGS-CADIKKVEAWQVLKHLRH 467

Qy        433 LTFHVG-GLPLRFDSSGNVDMEYDLKLW---VWQGSVPRLHDVGRFN-GSLRTERL---- 483
              | |   |  : ||  |::   | :  |      ||:   :|| :|  : : |||
Db        468 LNFTSNMGEQVTFDECGDLAGNYSIINWHLSPEDGSIV-FKEVGYYNVYAKKGERLFIND 526

Qy        484 -KIRWHTSDNQKPVSRCSRQCQEGQVRR-VKGFHSCCYDCVDCEAGSYRQNPDDIACTFC 541
               || |     : | | ||| |  |  :  ::|  :||::||:|  | |     |  ||  |
```

```
Db        527 EKILWSGFSREVPFSNCSRDCLAGTRKGIIEGEPTCCFECVECPDGEYSDETDASACDKC 586

Qy        542 GQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQA 601
              | || |  | |  :   ||:| ||  : | |   | : |    ||:|:  |::|:|:|
Db        587 PDDFWSNENHTSCIAKEIEFLSWTEPFGIALTLFAVLGIFLTAFVLGVFIKFRNTPIVKA 646

Qy        602 SGGPLA---CFGLVCLGLVCL-SVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAE 657
              :   |:    | |:|    |  | | | |:|    | :||   :    |:| : ::
Db        647 TNRELSYLLLFSLLC----CFSSSLFFIGEPQDWTCRLRQPAFGISFVLCISCILVKTNR 702

Qy        658 IFV--ESELPLS-----WADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTD 710
              : :  |:::| |     |  |         :|:| |   ::: :|  :|   ||   :
Db        703 VLLVFEAKIPTSFHRKWWGLNLQ--------FLLVFLCTFMQIVICAIWLNTAPPSSYRN 754

Qy        711 WHMLPTEAL-VHCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAML 769
              | |  | : : |   | :: |         || :||   |  |  |  :| |: :||:||
Db        755 -HELEDEIIFITCHEGSLMALGFLIGYTCLLAAICFFFAFKSRKLPENFNEAKFITFSML 813

Qy        770 AYFITWVSFVPLLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPE 827
              :|| |:||:|  |:       ||:: |:|    |:||     : |::: :|  || |
Db        814 IFFIVWISFIPAYASTYGKFVSAVEVIAILAASFGLLACIFFNKVYIILFKPSRNTIE 871
```

Search completed: December 16, 2005, 18:12:20
Job time : 53 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2005, 18:03:50 ; Search time 48 Seconds
(without alignments)
1467.493 Million cell updates/sec

Title: US-10-679-102-15
Perfect score: 4524
Sequence: 1 MLGPAVLGLSLWALLHPGTG.....GPGDAQGQNDGNTGNQGKHE 852

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 4524 | 100.0 | 852 | 2 | US-09-897-427A-6 | Sequence 6, Appli |
| 2 | 1153 | 25.5 | 841 | 2 | US-09-897-427A-2 | Sequence 2, Appli |
| 3 | 1089.5 | 24.1 | 1059 | 2 | US-09-134-513-2 | Sequence 2, Appli |
| 4 | 1068.5 | 23.6 | 1078 | 1 | US-08-485-588-7 | Sequence 7, Appli |
| 5 | 1068.5 | 23.6 | 1078 | 1 | US-08-484-565-7 | Sequence 7, Appli |
| 6 | 1068.5 | 23.6 | 1078 | 1 | US-08-480-751-7 | Sequence 7, Appli |
| 7 | 1068.5 | 23.6 | 1078 | 1 | US-08-943-986-7 | Sequence 7, Appli |
| 8 | 1068.5 | 23.6 | 1078 | 2 | US-08-353-784-7 | Sequence 7, Appli |
| 9 | 1068.5 | 23.6 | 1078 | 2 | US-08-484-719B-7 | Sequence 7, Appli |
| 10 | 1068.5 | 23.6 | 1078 | 2 | US-08-484-159-7 | Sequence 7, Appli |
| 11 | 1067.5 | 23.6 | 1078 | 2 | US-10-125-772-28 | Sequence 28, Appl |
| 12 | 1065 | 23.5 | 941 | 2 | US-10-125-772-8 | Sequence 8, Appli |
| 13 | 1064 | 23.5 | 941 | 2 | US-10-125-772-10 | Sequence 10, Appl |
| 14 | 1060 | 23.4 | 1085 | 1 | US-08-485-588-5 | Sequence 5, Appli |
| 15 | 1060 | 23.4 | 1085 | 1 | US-08-484-565-5 | Sequence 5, Appli |
| 16 | 1060 | 23.4 | 1085 | 1 | US-08-480-751-5 | Sequence 5, Appli |
| 17 | 1060 | 23.4 | 1085 | 1 | US-08-943-986-5 | Sequence 5, Appli |
| 18 | 1060 | 23.4 | 1085 | 2 | US-08-353-784-5 | Sequence 5, Appli |
| 19 | 1060 | 23.4 | 1085 | 2 | US-08-484-719B-5 | Sequence 5, Appli |
| 20 | 1060 | 23.4 | 1085 | 2 | US-08-484-159-5 | Sequence 5, Appli |
| 21 | 1057.5 | 23.4 | 1088 | 1 | US-08-485-588-6 | Sequence 6, Appli |
| 22 | 1057.5 | 23.4 | 1088 | 1 | US-08-484-565-6 | Sequence 6, Appli |
| 23 | 1057.5 | 23.4 | 1088 | 1 | US-08-480-751-6 | Sequence 6, Appli |
| 24 | 1057.5 | 23.4 | 1088 | 1 | US-08-943-986-6 | Sequence 6, Appli |
| 25 | 1057.5 | 23.4 | 1088 | 2 | US-08-353-784-6 | Sequence 6, Appli |
| 26 | 1057.5 | 23.4 | 1088 | 2 | US-08-484-719B-6 | Sequence 6, Appli |
| 27 | 1057.5 | 23.4 | 1088 | 2 | US-08-484-159-6 | Sequence 6, Appli |
| 28 | 1056 | 23.3 | 1027 | 2 | US-09-162-021B-2 | Sequence 2, Appli |
| 29 | 1056 | 23.3 | 1027 | 2 | US-08-268-051-8 | Sequence 8, Appli |
| 30 | 1056 | 23.3 | 1027 | 2 | US-10-125-772-2 | Sequence 2, Appli |
| 31 | 1054 | 23.3 | 1079 | 1 | US-08-485-588-8 | Sequence 8, Appli |
| 32 | 1054 | 23.3 | 1079 | 1 | US-08-484-565-8 | Sequence 8, Appli |
| 33 | 1054 | 23.3 | 1079 | 1 | US-08-480-751-8 | Sequence 8, Appli |
| 34 | 1054 | 23.3 | 1079 | 1 | US-08-943-986-8 | Sequence 8, Appli |
| 35 | 1054 | 23.3 | 1079 | 2 | US-08-353-784-8 | Sequence 8, Appli |
| 36 | 1054 | 23.3 | 1079 | 2 | US-08-484-719B-8 | Sequence 8, Appli |
| 37 | 1054 | 23.3 | 1079 | 2 | US-08-484-159-8 | Sequence 8, Appli |
| 38 | 1048.5 | 23.2 | 850 | 2 | US-10-125-772-12 | Sequence 12, Appl |
| 39 | 1014.5 | 22.4 | 839 | 2 | US-09-897-427A-4 | Sequence 4, Appli |
| 40 | 1009.5 | 22.3 | 877 | 2 | US-09-619-353-2 | Sequence 2, Appli |
| 41 | 1002.5 | 22.2 | 843 | 2 | US-09-361-631-1 | Sequence 1, Appli |
| 42 | 999.5 | 22.1 | 843 | 2 | US-09-361-631-2 | Sequence 2, Appli |
| 43 | 974 | 21.5 | 863 | 2 | US-09-619-353-14 | Sequence 14, Appl |
| 44 | 938.5 | 20.7 | 1219 | 1 | US-08-687-289A-6 | Sequence 6, Appli |
| 45 | 938.5 | 20.7 | 1219 | 2 | US-09-435-897-6 | Sequence 6, Appli |

ALIGNMENTS

RESULT 1
US-09-897-427A-6
; Sequence 6, Application US/09897427A
; Patent No. 6955887
; GENERAL INFORMATION:
; APPLICANT: ADLER, JON ELLIOT
; APPLICANT: LI, XIADONG
; APPLICANT: STAZEWSKI, LENA
; APPLICANT: XU, HONG
; APPLICANT: EHEVERRI, FERNANDO
; TITLE OF INVENTION: T1R HETERO-OLIGOMERIC TASTE RECEPTORS
; FILE REFERENCE: 078003-0282558
; CURRENT APPLICATION NUMBER: US/09/897,427A
; CURRENT FILING DATE: 2001-07-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 852
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-897-427A-6

Query Match 100.0%; Score 4524; DB 2; Length 852;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 852; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
```

```
Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420

Qy        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540

Qy        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600

Qy        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720

Qy        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780

Qy        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840

Qy        841 NDGNTGNQGKHE 852
              ||||||||||||
Db        841 NDGNTGNQGKHE 852
```

RESULT 2
US-09-897-427A-2
; Sequence 2, Application US/09897427A
; Patent No. 6955887
; GENERAL INFORMATION:
; APPLICANT: ADLER, JON ELLIOT
; APPLICANT: LI, XIADONG
; APPLICANT: STAZEWSKI, LENA
; APPLICANT: XU, HONG
; APPLICANT: EHEVERRI, FERNANDO
; TITLE OF INVENTION: T1R HETERO-OLIGOMERIC TASTE RECEPTORS
; FILE REFERENCE: 078003-0282558
; CURRENT APPLICATION NUMBER: US/09/897,427A
; CURRENT FILING DATE: 2001-07-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 841
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-897-427A-2

Query Match 25.5%; Score 1153; DB 2; Length 841;
Best Local Similarity 33.1%; Pred. No. 1.2e-101;
Matches 282; Conservative 149; Mismatches 360; Indels 62; Gaps 18;

```
Qy          6 VLGLSL-----WALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              ::|| |     ||     | :     |    : |||:| |||||        |: | ||
Db          8 LVGLQLLISCCWAFACHSTES----SPDFTLPGDYLLAGLFPL---HSGCLQVRHRPEVT 60

Qy         61 VCTR---FSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFL 117
              :| |   |: :|     ||:: |||||| : ||| : ||| |:| ||:     :  :|  |
Db         61 LCDRSCSFNEHGYHLFQAMRLGVEEINNSTALLPNITLGYQLYDVCSDS-ANVYATLRVL 119

Qy        118 AKAGSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSAR 177
              :  |   |    :   | | ||||||| |:  |  |    | ||:| :|| || | || :
```

```
Db        120 SLPGQHHIELQGDLLHYSPTVLAVIGPDSTNRAATTAALLSPFLVPMISYAASSETLSVK 179

Qy        178 ETFPSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICI 237
              :||| ||:|:|: |:     |||:||| |:: :|| |:||: |:     |  :||||
Db        180 RQYPSFLRTIPNDKYQVETMVLLLQKFGWTWISLVGSSDDYGQLGVQALENQATGQGICI 239

Qy        238 AHEGLVPL-PRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKV 296
              | : ::|   :  | |:   | ::  : |:   ||::|:|   |   |   : : |: ||
Db        240 AFKDIMPFSAQVGDERM---QCLMRHLAQAGATVVVVFSSRQLARVFFESVVLTNLTGKV 296

Qy        297 WVASEAWLTSDLVMGLPGMAQMGTVLGF-LQRGA--QLHEFPQYVKTHLALATDP----A 349
              |||||||  |  : |:||: ::| |||  :|: |   |  | :          |  |    :
Db        297 WVASEAWALSRHITGVPGIQRIGMVLGVAIQKRAVPGLKAFEEAYARADKKAPRPCHKGS 356

Qy        350 FCSALGEREQGLEEDVVGQRCPQCDCI---TLQNVSAGLNHHQTFSVYAAVYSVAQALHN 406
              :||:                | | :|      |:  : |  :    :: | |||:||  ||
Db        357 WCSS------------NQLCRECQAFMAHTMPKLKA-FSMSSAYNAYRAVYAVAHGLHQ 402

Qy        407 TLQCNASGCPAQDPVKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVP 466
              | | |||  ::  | |||||| :: : | :    : |: : :    |:: | | |
Db        403 LLGC-ASGACSRGRVYPWQLLEQIHKVHFLLHKDTVAFNDNRDPLSSYNIIAWDWNGPKW 461

Qy        467 RLHDVGRFNGS---LRTERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVD 523
              :|    |   |      ||:||  ||| | | ||  | ||  | | ||| ||::||
Db        462 TFTVLGSSTWSPVQLNINETKIQWHGKDNQVPKSVCSSDCLEGHQRVVTGFHHCCFECVP 521

Qy        524 CEAGSYRQNPDDIACTFCGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLV 583
              | ||::   |   |  ||::||:|| |  || |   |||  |    :||   :| | |:
Db        522 CGAGTFLNKSDLYRCQPCGKEEWAPEGSQTCFPRTVVFLALREHTSWVLLAANTLLLLLL 581

Qy        584 LAALGLFVHHRDSPLVQASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLP 643
              |   |||  | |:|:|:::|| |    |  |     |:  | |:|:   || :| |  |
Db        582 LGTAGLFAWHLDTPVVRSAGGRLCFLMLGSLAAGSGSLYGFFGEPTRPACLLRQALFALG 641

Qy        644 LTGCLSTLFLQAAEIFV----ESELPL---SWADRLSGCLRGPWAWLVVLLAMLVEVALC 696
              |  || | ::: :: :     :::|    :|       ::   | | |:::   :: :|
Db        642 FTIFLSCLTVRSFQLIIIFKFSTKVPTFYHAW-------VQNHGAGLFVMISSAAQLLIC 694

Qy        697 TWYLVAFPPEVVTDWHMLPTEALVHCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPG 756
              :|| : |    ::    |   :: |   : : | ||   |  |:   |  ::| : |
Db        695 LTWLVVWTPLPAREYQRFPHLVMLECTETNSLGFILAFLYNGLLSISAFACSYLGKDLPE 754

Qy        757 CYNRARGLTFAMLAYFITWVSFVPLLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYL 816
              || |: :||::|  |::|::|    :       ||  | | |  :       : ||:||:
Db        755 NYNEAKCVTFSLLFNFVSWIAFFTTASVYDGKYLPAANMMAGLSSLSSGFGGYFLPKCYV 814

Qy        817 LMRQPGLNTPEFF 829
              :: :| ||: | |
Db        815 ILCRPDLNSTEHF 827
```

RESULT 3
US-09-134-513-2
; Sequence 2, Application US/09134513
; Patent No. 6210964
; GENERAL INFORMATION:
; APPLICANT: Brown, Edward M.
; APPLICANT: Diaz, Ruben
; APPLICANT: Bai, Mei
; APPLICANT: Quinn, Stephen J.
; TITLE OF INVENTION: The Avian Extracellular Calcium-Sensing
; TITLE OF INVENTION: Receptor
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Vinson & Elkins L.L.P.
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.
; ZIP: 20004-1008
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/134,513
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sanzo, Michael A.
;      REGISTRATION NUMBER:  36,912
;      REFERENCE/DOCKET NUMBER:  BRI331/13003
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)639-6585
;      TELEFAX:  (202)639-6604
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1059 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
US-09-134-513-2

  Query Match             24.1%;  Score 1089.5;  DB 2;  Length 1059;
  Best Local Similarity   30.4%;  Pred. No. 2.3e-95;
  Matches  266;  Conservative  157;  Mismatches  350;  Indels  103;  Gaps   23;

Qy         26 SQQLRMKGDYVLGGLFPLG---EAEEAGLRSRTRPSSPVCTRFSSNGLLWALAMKMAVEE 82
              :|: : ||| :|||||:     |::  |:|  || |  | |::  |  |  ||  |:||
Db         23 NQRAQKKGDIILGGLFPIHFGVAAKDQDLKS--RPESVECIRYNFRGFRWLQAMIFAIEE 80

Qy         83 INNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLA--KAGSRDIAAYCNYTQYQPRVLA 140
              |||  :||| : ||| :||||:    |:: :| |:|  |  | ::  :|| ::: |  :|
Db         81 INNSPNLLPNMTLGYRIFDTCNTVSKALEATLSFVAQNKIDSLNLDEFCNCSEHIPSTIA 140

Qy        141 VIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETFPSFFRTVPSDRVQLTAAAEL 200
              |:|   | ::          | :|||||  :|  |||  :  | || ||:|:|  | || |::
Db        141 VVGATGSGVSTAVANLLGLFYIPQVSYASSSRLLSNKNQFKSFLRTIPNDEHQATAMADI 200

Qy        201 LQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHEGLVPLPRADDSRLGKVQDVL 260
              ::  | ||||  : :||:||| |:  |   |  | |||    |:    :  |   ::| |:
Db        201 IEYFRWNWVGTIAADDDYGRPGIEKFREEAEERDICIDFSELI----SQYSDEEEIQQVV 256

Qy        261 HQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVASEAWLTSDLVMGLPGMAQMGT 320
              :  |: :|:::|:|      |    :   :: |:|:|||| :| |:        :|:
Db        257 EVIQNSTARVIVVFSSGPDLEPLIKEIVRRNITGKIWLASEAWASSSLIAMPEFFRVIGS 316

Qy        321 VLGFLQRGAQLHEFPQYV-KTHLALATDPAFCSALGE---------------------- 356
              :||  :  |:  | ::: | |   : :  |     |
Db        317 TIGFALKAGQIPGFREFLQKVHPKKSANNGFAKEFWEETFNCYLPSESKNSPASASFHKA 376

Qy        357 REQGL--------------EEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVA 401
              |:||               :|::     |  |   |:          :::|| ||||:|
Db        377 HEEGLGAGNGTAAFRPPCTGDENITSVETPYMDFTHLR---------ISYNVYLAVYSIA 427

Qy        402 QALHNTLQCNA-------SGCPAQDPVKPWQLLENMYNLTFHVG-GLPLRFDSSGNVDME 453
              || :   |          |     |: ||:|::: :| |    |  : ||  |::
Db        428 HALQDIYTCTPGKGLFTNGSCADIKKVEAWQVLKHLRHLNFTSNMGEQVDFDEFGDLVGN 487

Qy        454 YDLKLW---VWQGSVPRLHDVGRFN-GSLRTERL-----KIRWHTSDNQKPVSRCSRQCQ 504
              | :  |       |||    :|| :|  : : |||     || |     : | | ||| |
Db        488 YSIINWHLSPEDGSVV-FEEVGHYNVYAKKGERLFINENKILWSGFSKEVPFSNCSRDCL 546

Qy        505 EGQVRR-VKGFHSCCYDCVDCEAGSYRQNPDDIACTFCGQDEWSPERSTRCFRRRSRFLA 563
              |  :  ::|  :||::||||  | |    |  ||  | :| || |  | |  ::  ||:
Db        547 PGTRKGIIEGEPTCCFECVDCPDGEYSDETDASACDKCPEDYWSNENHTSCIPKQIEFLS 606

Qy        564 WGEPAVLLLLLLSLALGLVLAALGLFVHHRDSPLVQASGGPLA---CFGLVCLGLVCL- 619
              | ||  :  | |   | : |     ||:|   |::|:|:|:   |:    | |:|
Db        607 WTEPPGIALTLFAVLGIFLTSFVLGVFTKFRNTPIVKATNRELSYLLLFSLLC----CFS 662

Qy        620 SVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV--ESELPLSWADRLSGCLR 677
              | | | |:|    |  :||   :    |:| : ::   : :  |:::| |        |

Db        663 SSLFFIGEPQNWTCRLRQPAFGISFVLCISCILVKTNRVLLVFEAKIPTS--------LH 714

Qy        678 GPW-----AWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEAL-VHCRTRSWVSFG 731
                    |       :|:| |    |:: :|  :|    ||     : | |   | : : |    | :: |
Db        715 RKWWGLNLQFLLVFLCTFVQIVICVIWLYTAPPSSYRN-HELEDEIIFITCHEGSLMALG 773

Qy        732 LAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVPLLANVQVVLRP 791
                      || :||    |  |  |  :| |: :||:|| :|| |:||:|   |:
Db        774 FLIGYTCLLAAICFFFAFKSRKLPENFNEAKFITFSMLIFFIVWISFIPAYASTYGKFVS 833

Qy        792 AVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPE 827
              ||:: |:|    |:||      : |::: :|  || |
Db        834 AVEVIAILAASFGLLACIFFNKVYIILFKPSRNTIE 869


RESULT 4
US-08-485-588-7
; Sequence 7, Application US/08485588
; Patent No. 5688938
;  GENERAL INFORMATION:
;    APPLICANT:  Edward M. Brown
;    APPLICANT:  Steven C. Hebert
;    APPLICANT:  Forrest H. Fuller
;    APPLICANT:  James E. Garrett, Jr.
;    TITLE OF INVENTION:  CALCIUM RECEPTOR-ACTIVE
;    TITLE OF INVENTION:  MOLECULES
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  First Interstate World Center
;      STREET:  Suite 4700
;      STREET:  633 West Fifth Street
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb storage
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FASTSEQ
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/485,588
;      FILING DATE:  7 June, 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;    PRIOR APPLICATION DATA:  including application
;    PRIOR APPLICATION DATA:  described below:  9
;      APPLICATION NUMBER:  08/353,784
;      FILING DATE:  9 December, 1994
;      APPLICATION NUMBER:  PCT/US/94/12117
;      FILING DATE:  21 October, 1994
;      APPLICATION NUMBER:  U.S. 08/292,827
;      FILING DATE:  23 August, 1994
;      APPLICATION NUMBER:  U.S. 08/141,248
;      FILING DATE:  22 October, 1993
;      APPLICATION NUMBER:  U.S. 08/009,389
;      FILING DATE:  23 February, 1993
;      APPLICATION NUMBER:  U.S. 08/017,127
;      FILING DATE:  12 February, 1993
;      APPLICATION NUMBER:  U.S. 07/934,161
;      FILING DATE:  21 August, 1992
;      APPLICATION NUMBER:  U.S. 07/834,044
;      FILING DATE:  11 February, 1992
;      APPLICATION NUMBER:  U.S. 07/749,451
;      FILING DATE:  23 August, 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Heber, Sheldon O.
;      REGISTRATION NUMBER:  38,179
;      REFERENCE/DOCKET NUMBER:  213/005
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
```

```
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1078 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-485-588-7

  Query Match            23.6%;  Score 1068.5;  DB 1;  Length 1078;
  Best Local Similarity  30.3%;  Pred. No. 2.5e-93;
  Matches  269;  Conservative  156;  Mismatches  366;  Indels   97;  Gaps   26;

Qy         12 WALL----HPGTGAPLCLSQQLRMKGDYVLGGLFPLG---EAEEAGLRSRTRPSSPVCTR 64
Db          8 WVLLALTWHTSAYGP---DQRAQKKGDIILGGLFPIHFGVAAKDQDLKS--RPESVECIR 62

Qy         65 FSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLA--KAGS 122
Db         63 YNFRGFRWLQAMIFAIEEINSSPALLPNLTLGYRIFDTCNTVSKALEATLSFVAQNKIDS 122

Qy        123 RDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETFPS 182
Db        123 LNLDEFCNCSEHIPSTIAVVGATGSGVSTAVANLLGLFYIPQVSYASSSRLLSNKNQFKS 182

Qy        183 FFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHEGL 242
Db        183 FLRTIPNDEHQATAMADIIEYFRWNWVGTIAADDDYGRPGIEKFREEAEERDICIDFSEL 242

Qy        243 VPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVASEA 302
Db        243 I----SQYSDEEEIQHVVEVIQNSTAKVIVVFSSGPDLEPLIKEIVRRNITGKIWLASEA 298

Qy        303 WLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYV-KTHLALATDPAFCSALGERE--- 358
Db        299 WASSSLIAMPQYFHVVGGTIGFALKAGQIPGFREFLKKVHPRKSVHNGFAKEFWEETFNC 358

Qy        359 ----------------QGLEEDVVGQRCPQCD------CITLQNVSA------GLNHHQ- 389
Db        359 HLQEGAKGPLPVDTFLRGHEES--GDRFSNSSTAFRPLCTGDENISSVETPYIDYTHLRI 416

Qy        390 TFSVYAAVYSVAQALHNTLQC--------NASGCPAQDPVKPWQLLENMYNLTF-HVGGL 440
Db        417 SYNVYLAVYSIAHALQDIYTCLPGRGLFTNGS-CADIKKVEAWQVLKHLRHLNFTNNMGE 475

Qy        441 PLRFDSSGNVDMEYDLKLW---VWQGSVPRLHDVGRFN-GSLRTERL-----KIRWHTSD 491
Db        476 QVTFDECGDLVGNYSIINWHLSPEDGSIV-FKEVGYYNVYAKKGERLFINEEKILWSGFS 534

Qy        492 NQKPVSRCSRQCQEGQVRR-VKGFHSCCYDCVDCEAGSYRQNPDDIACTFCGQDEWSPER 550
Db        535 REVPFSNCSRDCLAGTRKGIIEGEPTCCFECVECPDGEYSDETDASACNKCPDDFWSNEN 594

Qy        551 STRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQASGGPLA--- 607
Db        595 HTSCIAKEIEFLSWTEPFGIALTLFAVLGIFLTAFVLGVFIKFRNTPIVKATNRELSYLL 654

Qy        608 CFGLVCLGLVCL-SVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV--ESEL 664
Db        655 LFSLLC----CFSSSLFFIGEPQDWTCRLRQPAFGISFVLCISCILVKTNRVLLVFEAKI 710

Qy        665 PLS-----WADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEAL 719
Db        711 PTSFHRKWWGLNLQ--------FLLVFLCTFMQIVICVIWLYTAPPSSYRNQELEDEIIF 762

Qy        720 VHCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFV 779
Db        763 ITCHEGSLMALGFLIGYTCLLAAICFFFAFKSRKLPENFNEAKFITFSMLIFFIVWISFI 822

Qy        780 PLLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPE 827
Db        823 PAYASTYGKFVSAVEVIAILAASFGLLACIFFNKIYIILFKPSRNTIE 870

RESULT 5
US-08-484-565-7
; Sequence 7, Application US/08484565
; Patent No. 5763569
;  GENERAL INFORMATION:
;    APPLICANT:  Edward M. Brown
;    APPLICANT:  Steven C. Hebert
;    APPLICANT:  James E. Garrett, Jr.
;    TITLE OF INVENTION:  CALCIUM RECEPTOR-ACTIVE
;    TITLE OF INVENTION:  MOLECULES
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  First Interstate World Center
;      STREET:  Suite 4700
;      STREET:  633 West Fifth Street
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb storage
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FASTSEQ
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,565
;      FILING DATE:  7 June, 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;    PRIOR APPLICATION DATA:  including application
;    PRIOR APPLICATION DATA:  described below:  9
;      APPLICATION NUMBER:  08/353,784
;      FILING DATE:  9 December, 1994
;      APPLICATION NUMBER:  PCT/US/94/12117
;      FILING DATE:  21 October, 1994
;      APPLICATION NUMBER:  U.S. 08/292,827
;      FILING DATE:  23 August, 1994
;      APPLICATION NUMBER:  U.S. 08/141,248
;      FILING DATE:  22 October, 1993
;      APPLICATION NUMBER:  U.S. 08/009,389
;      FILING DATE:  23 February, 1993
;      APPLICATION NUMBER:  U.S. 08/017,127
;      FILING DATE:  12 February, 1993
;      APPLICATION NUMBER:  U.S. 07/934,161
;      FILING DATE:  21 August, 1992
;      APPLICATION NUMBER:  U.S. 07/834,044
;      FILING DATE:  11 February, 1992
;      APPLICATION NUMBER:  U.S. 07/749,451
;      FILING DATE:  23 August, 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Heber, Sheldon O.
;      REGISTRATION NUMBER:  38,179
;      REFERENCE/DOCKET NUMBER:  213/006
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1078 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-484-565-7

  Query Match            23.6%;  Score 1068.5;  DB 1;  Length 1078;
  Best Local Similarity  30.3%;  Pred. No. 2.5e-93;
  Matches  269;  Conservative  156;  Mismatches  366;  Indels   97;  Gaps   26;

Qy         12 WALL----HPGTGAPLCLSQQLRMKGDYVLGGLFPLG---EAEEAGLRSRTRPSSPVCTR 64
```

```
Db          8 WVLLALTWHTSAYGP---DQRAQKKGDIILGGLFPIHFGVAAKDQDLKS--RPESVECIR 62

Qy         65 FSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLA--KAGS 122
              ::  |  |  ||  |:||||:   ||| | ||| :||||:    |:: :| |:|  |  |
Db         63 YNFRGFRWLQAMIFAIEEINSSPALLPNLTLGYRIFDTCNTVSKALEATLSFVAQNKIDS 122

Qy        123 RDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETFPS 182
              ::  :|| ::: |  :||:|   | ::          | :||||| :|  ||| :  | |
Db        123 LNLDEFCNCSEHIPSTIAVVGATGSGVSTAVANLLGLFYIPQVSYASSSRLLSNKNQFKS 182

Qy        183 FFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHEGL 242
              | ||:|:|  | || |:::: | ||||  : :||:||| |:  |    |  | |||    |
Db        183 FLRTIPNDEHQATAMADIIEYFRWNWVGTIAADDDYGRPGIEKFREEAEERDICIDFSEL 242

Qy        243 VPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVASEA 302
              :   :  |   ::| |:  :  |: :|:::|:|      |    :   :: |:|:|||| 
Db        243 I----SQYSDEEEIQHVVEVIQNSTAKVIVVFSSGPDLEPLIKEIVRRNITGKIWLASEA 298

Qy        303 WLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYV-KTHLALATDPAFCSALGERE--- 358
              | :| |:        :|  :||  :  |:  | ::: | |   :    |      |
Db        299 WASSSLIAMPQYFHVVGGTIGFALKAGQIPGFREFLKKVHPRKSVHNGFAKEFWEETFNC 358

Qy        359 ----------------QGLEEDVVGQRCPQCD------CITLQNVSA------GLNHHQ- 389
                           :| ||   | |           |   :|:|:         | :
Db        359 HLQEGAKGPLPVDTFLRGHEES--GDRFSNSSTAFRPLCTGDENISSVETPYIDYTHLRI 416

Qy        390 TFSVYAAVYSVAQALHNTLQC--------NASGCPAQDPVKPWQLLENMYNLTF-HVGGL 440
              :::|| ||||:| || :   |         | | |     |: ||:|::: :| | :  |
Db        417 SYNVYLAVYSIAHALQDIYTCLPGRGLFTNGS-CADIKKVEAWQVLKHLRHLNFTNNMGE 475

Qy        441 PLRFDSSGNVDMEYDLKLW---VWQGSVPRLHDVGRFN-GSLRTERL-----KIRWHTSD 491
              : ||  |::   | :  |      ||:    :|| :|  : : |||      || |
Db        476 QVTFDECGDLVGNYSIINWHLSPEDGSIV-FKEVGYYNVYAKKGERLFINEEKILWSGFS 534

Qy        492 NQKPVSRCSRQCQEGQVRR-VKGFHSCCYDCVDCEAGSYRQNPDDIACTFCGQDEWSPER 550
              : | | ||| |  |  :  ::|  :||::||:|  | |     |  ||  |  | || |
Db        535 REVPFSNCSRDCLAGTRKGIIEGEPTCCFECVECPDGEYSDETDASACNKCPDDFWSNEN 594

Qy        551 STRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQASGGPLA--- 607
              | |  :   ||:| ||  : | |   | : |    ||:|:  |::|:|:|:   |:
Db        595 HTSCIAKEIEFLSWTEPFGIALTLFAVLGIFLTAFVLGVFIKFRNTPIVKATNRELSYLL 654

Qy        608 CFGLVCLGLVCL-SVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV--ESEL 664
              | |:|    | | | | |:|    |  :||   :    |:| : ::   : :  |:::
Db        655 LFSLLC----CFSSSLFFIGEPQDWTCRLRQPAFGISFVLCISCILVKTNRVLLVFEAKI 710

Qy        665 PLS-----WADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEAL 719
              | |     |  |          :|:| |  ::: :|  :|   ||    :  :
Db        711 PTSFHRKWWGLNLQ--------FLLVFLCTFMQIVICVIWLYTAPPSSYRNQELEDEIIF 762

Qy        720 VHCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFV 779
              : |   | :: |        || :||   |  |  |  :| |: :||:|| :|| |:||:
Db        763 ITCHEGSLMALGFLIGYTCLLAAICFFFAFKSRKLPENFNEAKFITFSMLIFFIVWISFI 822

Qy        780 PLLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPE 827
              | |:       ||:: |:|    |:||     : |::: :|  || |
Db        823 PAYASTYGKFVSAVEVIAILAASFGLLACIFFNKIYIILFKPSRNTIE 870


RESULT 6
US-08-480-751-7
; Sequence 7, Application US/08480751
; Patent No. 5858684
;  GENERAL INFORMATION:
;    APPLICANT:  Edward F. Nemeth
;    APPLICANT:  Edward M. Brown
;    APPLICANT:  Steven C. Hebert
;    APPLICANT:  Forrest H. Fuller
;    APPLICANT:  James E. Garrett, Jr.
;    TITLE OF INVENTION:  CALCIUM RECEPTOR-ACTIVE
;    TITLE OF INVENTION:  MOLECULES
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  First Interstate World Center
;      STREET:  Suite 4700
;      STREET:  633 West Fifth Street
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb storage
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FASTSEQ
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/480,751
;      FILING DATE:  7 June, 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;    PRIOR APPLICATION DATA:  including application
;    PRIOR APPLICATION DATA:  described below:  9
;      APPLICATION NUMBER:  08/353,784
;      FILING DATE:  9 December, 1994
;      APPLICATION NUMBER:  PCT/US/94/12117
;      FILING DATE:  21 October, 1994
;      APPLICATION NUMBER:  U.S. 08/292,827
;      FILING DATE:  23 August, 1994
;      APPLICATION NUMBER:  U.S. 08/141,248
;      FILING DATE:  22 October, 1993
;      APPLICATION NUMBER:  U.S. 08/009,389
;      FILING DATE:  23 February, 1993
;      APPLICATION NUMBER:  U.S. 08/017,127
;      FILING DATE:  12 February, 1993
;      APPLICATION NUMBER:  U.S. 07/934,161
;      FILING DATE:  21 August, 1992
;      APPLICATION NUMBER:  U.S. 07/834,044
;      FILING DATE:  11 February, 1992
;      APPLICATION NUMBER:  U.S. 07/749,451
;      FILING DATE:  23 August, 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Heber, Sheldon O.
;      REGISTRATION NUMBER:  38,179
;      REFERENCE/DOCKET NUMBER:  213/004
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1078 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-480-751-7

  Query Match           23.6%;  Score 1068.5;  DB 1;  Length 1078;
  Best Local Similarity  30.3%;  Pred. No. 2.5e-93;
  Matches  269;  Conservative  156;  Mismatches  366;  Indels   97;  Gaps   26;

Qy         12 WALL----HPGTGAPLCLSQQLRMKGDYVLGGLFPLG---EAEEAGLRSRTRPSSPVCTR 64
              | ||    |     |    |: : ||| :||||||:     |::  |:|  || |  | |
Db          8 WVLLALTWHTSAYGP---DQRAQKKGDIILGGLFPIHFGVAAKDQDLKS--RPESVECIR 62

Qy         65 FSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLA--KAGS 122
              ::  |  |  ||  |:||||:   ||| | ||| :||||:    |:: :| |:|  |  |
Db         63 YNFRGFRWLQAMIFAIEEINSSPALLPNLTLGYRIFDTCNTVSKALEATLSFVAQNKIDS 122

Qy        123 RDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETFPS 182
              ::  :|| ::: |  :||:|   | ::          | :||||| :|  ||| :  | |
Db        123 LNLDEFCNCSEHIPSTIAVVGATGSGVSTAVANLLGLFYIPQVSYASSSRLLSNKNQFKS 182

Qy        183 FFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHEGL 242
              | ||:|:|  | || |:::: | ||||  : :||:||| |:  |    |  | |||    |
Db        183 FLRTIPNDEHQATAMADIIEYFRWNWVGTIAADDDYGRPGIEKFREEAEERDICIDFSEL 242
```

```
Qy        243 VPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVASEA 302
              :    :  |   ::| |:  :  |: :|:::|:|      |    :   :: |:|:|||| 
Db        243 I----SQYSDEEEIQHVVEVIQNSTAKVIVVFSSGPDLEPLIKEIVRRNITGKIWLASEA 298

Qy        303 WLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYV-KTHLALATDPAFCSALGERE--- 358
              | :| |:        :|  :||  :  |:  | ::: | |   :    |      |
Db        299 WASSSLIAMPQYFHVVGGTIGFALKAGQIPGFREFLKKVHPRKSVHNGFAKEFWEETFNC 358

Qy        359 ----------------QGLEEDVVGQRCPQCD------CITLQNVSA------GLNHHQ- 389
                             :| ||   | |           |   :|:|:          | :
Db        359 HLQEGAKGPLPVDTFLRGHEES--GDRFSNSSTAFRPLCTGDENISSVETPYIDYTHLRI 416

Qy        390 TFSVYAAVYSVAQALHNTLQC--------NASGCPAQDPVKPWQLLENMYNLTF-HVGGL 440
              :::|| ||||:| || :   |        | | |      |: ||:|::: :| | :  |
Db        417 SYNVYLAVYSIAHALQDIYTCLPGRGLFTNGS-CADIKKVEAWQVLKHLRHLNFTNNMGE 475

Qy        441 PLRFDSSGNVDMEYDLKLW---VWQGSVPRLHDVGRFN-GSLRTERL-----KIRWHTSD 491
               : ||  |::   | :  |      ||:    :|| :|  : : |||     || |
Db        476 QVTFDECGDLVGNYSIINWHLSPEDGSIV-FKEVGYYNVYAKKGERLFINEEKILWSGFS 534

Qy        492 NQKPVSRCSRQCQEGQVRR-VKGFHSCCYDCVDCEAGSYRQNPDDIACTFCGQDEWSPER 550
               : | | ||| |  |  :  ::|  :||::||:|  | |     |  ||  |  | || |
Db        535 REVPFSNCSRDCLAGTRKGIIEGEPTCCFECVECPDGEYSDETDASACNKCPDDFWSNEN 594

Qy        551 STRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQASGGPLA--- 607
              | |  :   ||:| ||  : | |   | : |    ||:|:  |::|:|:|:   |:
Db        595 HTSCIAKEIEFLSWTEPFGIALTLFAVLGIFLTAFVLGVFIKFRNTPIVKATNRELSYLL 654

Qy        608 CFGLVCLGLVCL-SVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV--ESEL 664
              | |:|    |  | | | |:|    |  :||   :    |:| : ::   : :  |:::
Db        655 LFSLLC----CFSSSLFFIGEPQDWTCRLRQPAFGISFVLCISCILVKTNRVLLVFEAKI 710

Qy        665 PLS-----WADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEAL 719
              | |     |  |          :|:| |   ::: :|  :|   ||     :  :
Db        711 PTSFHRKWWGLNLQ--------FLLVFLCTFMQIVICVIWLYTAPPSSYRNQELEDEIIF 762

Qy        720 VHCRTRSWVSPGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFV 779
              : |   | :: |          || :||   |  |  | :| |: :||:|| :|| |:||:
Db        763 ITCHEGSLMALGFLIGYTCLLAAICFFFAFKSRKLPENFNEAKFITFSMLIFFIVWISFI 822

Qy        780 PLLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPE 827
              |  |:       ||:: |:|    |:||     : |::: :|  || |
Db        823 PAYASTYGKFVSAVEVIAILAASFGLLACIFFNKIYIILFKPSRNTIE 870
```

```
RESULT 7
US-08-943-986-7
; Sequence 7, Application US/08943986
; Patent No. 5962314
;  GENERAL INFORMATION:
;    APPLICANT:  Edward M. Brown
;    APPLICANT:  Steven C. Hebert
;    APPLICANT:  James E. Garrett, Jr.
;    TITLE OF INVENTION:  CALCIUM RECEPTOR-ACTIVE
;    TITLE OF INVENTION:  MOLECULES
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  First Interstate World Center
;      STREET:  Suite 4700
;      STREET:  633 West Fifth Street
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb storage
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FASTSEQ
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/943,986
;      FILING DATE:  03-OCT-1997
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/484,565
;      FILING DATE:  7-June-1995
;      APPLICATION NUMBER:  08/353,784
;      FILING DATE:  9 December, 1994
;      APPLICATION NUMBER:  PCT/US/94/12117
;      FILING DATE:  21 October, 1994
;      APPLICATION NUMBER:  U.S. 08/292,827
;      FILING DATE:  23 August, 1994
;      APPLICATION NUMBER:  U.S. 08/141,248
;      FILING DATE:  22 October, 1993
;      APPLICATION NUMBER:  U.S. 08/009,389
;      FILING DATE:  23 February, 1993
;      APPLICATION NUMBER:  U.S. 08/017,127
;      FILING DATE:  12 February, 1993
;      APPLICATION NUMBER:  U.S. 07/934,161
;      FILING DATE:  21 August, 1992
;      APPLICATION NUMBER:  U.S. 07/834,044
;      FILING DATE:  11 February, 1992
;      APPLICATION NUMBER:  U.S. 07/749,451
;      FILING DATE:  23 August, 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Heber, Sheldon O.
;      REGISTRATION NUMBER:  38,179
;      REFERENCE/DOCKET NUMBER:  213/006
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1078 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-943-986-7

  Query Match           23.6%;  Score 1068.5;  DB 1;  Length 1078;
  Best Local Similarity  30.3%;  Pred. No. 2.5e-93;
  Matches  269;  Conservative  156;  Mismatches  366;  Indels   97;  Gaps   26;

Qy         12 WALL----HPGTGAPLCLSQQLRMKGDYVLGGLFPLG---EAEEAGLRSRTRPSSPVCTR 64
              | ||    |     |    |: : ||| :||||||:      |::  |:|  || |  | |
Db          8 WVLLALTWHTSAYGP---DQRAQKKGDIILGGLFPIHFGVAAKDQDLKS--RPESVECIR 62

Qy         65 FSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLA--KAGS 122
              ::  |  |  ||  |:|||:   ||| | ||| :||||:    |:: :| |:|  |  |
Db         63 YNFRGFRWLQAMIFAIEEINSSPALLPNLTLGYRIFDTCNTVSKALEATLSFVAQNKIDS 122

Qy        123 RDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETFPS 182
               ::  :|| ::: |  :||:|    | ::          | :|||| :|  ||| :  | |
Db        123 LNLDEFCNCSEHIPSTIAVVGATGSGVSTAVANLLGLFYIPQVSYASSSRLLSNKNQFKS 182

Qy        183 FFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHEGL 242
              | ||:|:|  | || |:::: | ||||  : :||:||| |:  |    |  | |||     |
Db        183 FLRTIPNDEHQATAMADIIEYFRWNWVGTIAADDDYGRPGIEKFREEAEERDICIDFSEL 242

Qy        243 VPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVASEA 302
              :    :  |   ::| |:  :  |: :|:::|:|      |    :   :: |:|:|||| 
Db        243 I----SQYSDEEEIQHVVEVIQNSTAKVIVVFSSGPDLEPLIKEIVRRNITGKIWLASEA 298

Qy        303 WLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYV-KTHLALATDPAFCSALGERE--- 358
              | :| |:        :|  :||  :  |:  | ::: | |   :    |      |
Db        299 WASSSLIAMPQYFHVVGGTIGFALKAGQIPGFREFLKKVHPRKSVHNGFAKEFWEETFNC 358

Qy        359 ----------------QGLEEDVVGQRCPQCD------CITLQNVSA------GLNHHQ- 389
                             :| ||   | |           |   :|:|:          | :
Db        359 HLQEGAKGPLPVDTFLRGHEES--GDRFSNSSTAFRPLCTGDENISSVETPYIDYTHLRI 416

Qy        390 TFSVYAAVYSVAQALHNTLQC--------NASGCPAQDPVKPWQLLENMYNLTF-HVGGL 440
              :::|| ||||:| || :   |        | | |      |: ||:|::: :| | :  |
```

```
Db        417 SYNVYLAVYSIAHALQDIYTCLPGRGLFTNGS-CADIKKVEAWQVLKHLRHLNFTNNMGE 475

Qy        441 PLRFDSSGNVDMEYDLKLW---VWQGSVPRLHDVGRFN-GSLRTERL-----KIRWHTSD 491

Db        476 QVTFDECGDLVGNYSIINWHLSPEDGSIV-FKEVGYYNVYAKKGERLFINEEKILWSGFS 534

Qy        492 NQKPVSRCSRQCQEGQVRR-VKGFHSCCYDCVDCEAGSYRQNPDDIACTFCGQDEWSPER 550

Db        535 REVPFSNCSRDCLAGTRKGIIEGEPTCCFECVECPDGEYSDETDASACNKCPDDFWSNEN 594

Qy        551 STRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQASGGPLA--- 607

Db        595 HTSCIAKEIEFLSWTEPFGIALTLFAVLGIFLTAFVLGVFIKFRNTPIVKATNRELSYLL 654

Qy        608 CFGLVCLGLVCL-SVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV--ESEL 664

Db        655 LFSLLC----CFSSSLFFIGEPQDWTCRLRQPAFGISFVLCISCILVKTNRVLLVFEAKI 710

Qy        665 PLS-----WADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEAL 719

Db        711 PTSFHRKWWGLNLQ--------FLLVFLCTFMQIVICVIWLYTAPPSSYRNQELEDEIIF 762

Qy        720 VHCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFV 779

Db        763 ITCHEGSLMALGFLIGYTCLLAAICFFFAFKSRKLPENFNEAKFITFSMLIFFIVWISFI 822

Qy        780 PLLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPE 827

Db        823 PAYASTYGKFVSAVEVIAILAASFGLLACIFFNKIYIILFKPSRNTIE 870


RESULT 8
US-08-353-784-7
; Sequence 7, Application US/08353784
; Patent No. 6011068
;  GENERAL INFORMATION:
;    APPLICANT:  Edward F. Nemeth, Edward M.
;    APPLICANT:  Brown, Steven C. Hebert,
;    APPLICANT:  Bradford C. Van Wagenen, Manuel
;    APPLICANT:  F. Balandrin, Forrest H. Fuller,
;    APPLICANT:  Eric G. DelMar, and Scott T. Moe
;    TITLE OF INVENTION:  CALCIUM RECEPTOR-ACTIVE
;    TITLE OF INVENTION:  MOLECULES
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  First Interstate World Center
;      STREET:  Suite 4700
;      STREET:  633 West Fifth Street
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb storage
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FASTSEQ
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/353,784
;      FILING DATE:  9 December, 1994
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;    PRIOR APPLICATION DATA:  including application
;    PRIOR APPLICATION DATA:  described below:  8
;      APPLICATION NUMBER:  PCT/US/94/12117
;      FILING DATE:  21 October, 1994
;      APPLICATION NUMBER:  U.S. 08/292,827
;      FILING DATE:  23 August, 1994
;      APPLICATION NUMBER:  U.S. 08/141,248
;      FILING DATE:  22 October, 1993
;      APPLICATION NUMBER:  U.S. 08/009,389
;      FILING DATE:  23 February, 1993
;      APPLICATION NUMBER:  U.S. 08/017,127
;      FILING DATE:  12 February, 1993
;      APPLICATION NUMBER:  U.S. 07/934,161
;      FILING DATE:  21 August, 1992
;      APPLICATION NUMBER:  U.S. 07/834,044
;      FILING DATE:  11 February, 1992
;      APPLICATION NUMBER:  U.S. 07/749,451
;      FILING DATE:  23 August, 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Heber, Sheldon O.
;      REGISTRATION NUMBER:  38,179
;      REFERENCE/DOCKET NUMBER:  209/069
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1078 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-353-784-7

  Query Match             23.6%;  Score 1068.5;  DB 2;  Length 1078;
  Best Local Similarity   30.3%;  Pred. No. 2.5e-93;
  Matches  269;  Conservative  156;  Mismatches  366;  Indels   97;  Gaps   26;

Qy         12 WALL----HPGTGAPLCLSQQLRMKGDYVLGGLFPLG---EAEEAGLRSRTRPSSPVCTR 64

Db          8 WVLLALTWHTSAYGP---DQRAQKKGDIILGGLFPIHFGVAAKDQDLKS--RPESVECIR 62

Qy         65 FSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLA--KAGS 122

Db         63 YNFRGFRWLQAMIFAIEEINSSPALLPNLTLGYRIFDTCNTVSKALEATLSFVAQNKIDS 122

Qy        123 RDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETFPS 182

Db        123 LNLDEFCNCSEHIPSTIAVVGATGSGVSTAVANLLGLFYIPQVSYASSSRLLSNKNQFKS 182

Qy        183 FFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHEGL 242

Db        183 FLRTIPNDEHQATAMADIIEYFRWNWVGTIAADDDYGRPGIEKFREEAEERDICIDFSEL 242

Qy        243 VPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVASEA 302

Db        243 I----SQYSDEEEIQHVVEVIQNSTAKVIVVFSSGPDLEPLIKEIVRRNITGKIWLASEA 298

Qy        303 WLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYV-KTHLALATDPAFCSALGERE--- 358

Db        299 WASSSLIAMPQYFHVVGGTIGFALKAGQIPGFREFLKKVHPRKSVHNGFAKEFWEETFNC 358

Qy        359 ----------------QGLEEDVVGQRCPQCD------CITLQNVSA------GLNHHQ- 389

Db        359 HLQEGAKGPLPVDTFLRGHEES--GDRFSNSSTAFRPLCTGDENISSVETPYIDYTHLRI 416

Qy        390 TFSVYAAVYSVAQALHNTLQC--------NASGCPAQDPVKPWQLLENMYNLTF-HVGGL 440

Db        417 SYNVYLAVYSIAHALQDIYTCLPGRGLFTNGS-CADIKKVEAWQVLKHLRHLNFTNNMGE 475

Qy        441 PLRFDSSGNVDMEYDLKLW---VWQGSVPRLHDVGRFN-GSLRTERL-----KIRWHTSD 491

Db        476 QVTFDECGDLVGNYSIINWHLSPEDGSIV-FKEVGYYNVYAKKGERLFINEEKILWSGFS 534

Qy        492 NQKPVSRCSRQCQEGQVRR-VKGFHSCCYDCVDCEAGSYRQNPDDIACTFCGQDEWSPER 550

Db        535 REVPFSNCSRDCLAGTRKGIIEGEPTCCFECVECPDGEYSDETDASACNKCPDDFWSNEN 594

Qy        551 STRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQASGGPLA--- 607

Db        595 HTSCIAKEIEFLSWTEPFGIALTLFAVLGIFLTAFVLGVFIKFRNTPIVKATNRELSYLL 654

Qy        608 CFGLVCLGLVCL-SVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV--ESEL 664
```

```
Db        655 LFSLLC----CFSSSLFFIGEPQDWTCRLRQPAFGISFVLCISCILVKTNRVLLVFEAKI 710

Qy        665 PLS-----WADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEAL 719

Db        711 PTSFHRKWWGLNLQ--------FLLVFLCTFMQIVICVIWLYTAPPSSYRNQELEDEIIF 762

Qy        720 VHCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFV 779

Db        763 ITCHEGSLMALGFLIGYTCLLAAICFFFAFKSRKLPENFNEAKFITFSMLIFFIVWISFI 822

Qy        780 PLLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPE 827

Db        823 PAYASTYGKFVSAVEVIAILAASFGLLACIFFNKIYIILFKPSRNTIE 870


RESULT 9
US-08-484-719B-7
; Sequence 7, Application US/08484719B
; Patent No. 6031003
;  GENERAL INFORMATION:
;    APPLICANT:  Edward F. Nemeth, Edward M.
;    APPLICANT:  Brown, Steven C. Hebert,
;    APPLICANT:  Bradford C. Van Wagenen,
;    APPLICANT:  Manuel F. Balandrin,
;    APPLICANT:  Forrest H. Fuller, Eric G.
;    APPLICANT:  Delmar, Scott T. Moe
;    TITLE OF INVENTION:  CALCIUM RECEPTOR-ACTIVE
;    TITLE OF INVENTION:  MOLECULES
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  First Interstate World Center
;      STREET:  Suite 4700
;      STREET:  633 West Fifth Street
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb storage
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  MS Word
;      SOFTWARE:  FastSEQ for Windows Version 3.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,719B
;      FILING DATE:  7 June, 1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/353,784
;      FILING DATE:  9 December, 1994
;      APPLICATION NUMBER:  PCT/US/94/12117
;      FILING DATE:  21 October, 1994
;      APPLICATION NUMBER:  U.S. 08/292,827
;      FILING DATE:  23 August, 1994
;      APPLICATION NUMBER:  U.S. 08/141,248
;      FILING DATE:  22 October, 1993
;      APPLICATION NUMBER:  U.S. 08/009,389
;      FILING DATE:  23 February, 1993
;      APPLICATION NUMBER:  U.S. 08/017,127
;      FILING DATE:  12 February, 1993
;      APPLICATION NUMBER:  U.S. 07/934,161
;      FILING DATE:  21 August, 1992
;      APPLICATION NUMBER:  U.S. 07/834,044
;      FILING DATE:  11 February, 1992
;      APPLICATION NUMBER:  U.S. 07/749,451
;      FILING DATE:  23 August, 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Douglas C. Murdock
;      REGISTRATION NUMBER:  37,549
;      REFERENCE/DOCKET NUMBER:  213/007
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1078 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-484-719B-7

  Query Match             23.6%;  Score 1068.5;  DB 2;  Length 1078;
  Best Local Similarity   30.3%;  Pred. No. 2.5e-93;
  Matches  269;  Conservative  156;  Mismatches  366;  Indels   97;  Gaps   26;

Qy         12 WALL----HPGTGAPLCLSQQLRMKGDYVLGGLFPLG---EAEEAGLRSRTRPSSPVCTR 64

Db          8 WVLLALTWHTSAYGP---DQRAQKKGDIILGGLFPIHFGVAAKDQDLKS--RPESVECIR 62

Qy         65 FSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLA--KAGS 122

Db         63 YNFRGFRWLQAMIFAIEEINSSPALLPNLTLGYRIFDTCNTVSKALEATLSFVAQNKIDS 122

Qy        123 RDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETFPS 182

Db        123 LNLDEFCNCSEHIPSTIAVVGATGSGVSTAVANLLGLFYIPQVSYASSSRLLSNKNQFKS 182

Qy        183 FFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHEGL 242

Db        183 FLRTIPNDEHQATAMADIIEYFRWNWVGTIAADDDYGRPGIEKFREEAEERDICIDFSEL 242

Qy        243 VPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVASEA 302

Db        243 I----SQYSDEEEIQHVVEVIQNSTAKVIVVFSSGPDLEPLIKEIVRRNITGKIWLASEA 298

Qy        303 WLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYV-KTHLALATDPAFCSALGERE--- 358

Db        299 WASSSLIAMPQYFHVVGGTIGFALKAGQIPGFREFLKKVHPRKSVHNGFAKEFWEETFNC 358

Qy        359 ----------------QGLEEDVVGQRCPQCD------CITLQNVSA------GLNHHQ- 389

Db        359 HLQEGAKGPLPVDTFLRGHEES--GDRFSNSSTAFRPLCTGDENISSVETPYIDYTHLRI 416

Qy        390 TFSVYAAVYSVAQALHNTLQC--------NASGCPAQDPVKPWQLLENMYNLTF-HVGGL 440

Db        417 SYNVYLAVYSIAHALQDIYTCLPGRGLFTNGS-CADIKKVEAWQVLKHLRHLNFTNNMGE 475

Qy        441 PLRFDSSGNVDMEYDLKLW---VWQGSVPRLHDVGRFN-GSLRTERL-----KIRWHTSD 491

Db        476 QVTFDECGDLVGNYSIINWHLSPEDGSIV-FKEVGYYNVYAKKGERLFINEEKILWSGFS 534

Qy        492 NQKPVSRCSRQCQEGQVRR-VKGFHSCCYDCVDCEAGSYRQNPDDIACTFCGQDEWSPER 550

Db        535 REVPFSNCSRDCLAGTRKGIIEGEPTCCFECVECPDGEYSDETDASACNKCPDDFWSNEN 594

Qy        551 STRCFRRRSRFLAWGEPAVLLLLLLSLALGLVLAALGLFVHHRDSPLVQASGGPLA--- 607

Db        595 HTSCIAKEIEFLSWTEPFGIALTLFAVLGIFLTAFVLGVFIKFRNTPIVKATNRELSYLL 654

Qy        608 CFGLVCLGLVCL-SVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV--ESEL 664

Db        655 LFSLLC----CFSSSLFFIGEPQDWTCRLRQPAFGISFVLCISCILVKTNRVLLVFEAKI 710

Qy        665 PLS-----WADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEAL 719

Db        711 PTSFHRKWWGLNLQ--------FLLVFLCTFMQIVICVIWLYTAPPSSYRNQELEDEIIF 762

Qy        720 VHCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFV 779

Db        763 ITCHEGSLMALGFLIGYTCLLAAICFFFAFKSRKLPENFNEAKFITFSMLIFFIVWISFI 822

Qy        780 PLLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPE 827
```

```
Db        823 PAYASTYGKFVSAVEVIAILAASFGLLACIFFNKIYIILFKPSRNTIE 870

RESULT 10
US-08-484-159-7
; Sequence 7, Application US/08484159
; Patent No. 6313146
;  GENERAL INFORMATION:
;    APPLICANT:  Bradford C. Van Wagenen
;    APPLICANT:  Manuel F. Balandrin
;    APPLICANT:  Eric G. Del Mar
;    APPLICANT:  Edward F. Nemeth
;    TITLE OF INVENTION:  CALCIUM RECEPTOR-ACTIVE
;    TITLE OF INVENTION:  MOLECULES
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  First Interstate World Center
;      STREET:  Suite 4700
;      STREET:  633 West Fifth Street
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb storage
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FASTSEQ
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,159
;      FILING DATE:  7 June, 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;    PRIOR APPLICATION DATA:  including application
;    PRIOR APPLICATION DATA:  described below:  9
;      APPLICATION NUMBER:  08/353,784
;      FILING DATE:  9 December, 1994
;      APPLICATION NUMBER:  PCT/US/94/12117
;      FILING DATE:  21 October, 1994
;      APPLICATION NUMBER:  U.S. 08/292,827
;      FILING DATE:  23 August, 1994
;      APPLICATION NUMBER:  U.S. 08/141,248
;      FILING DATE:  22 October, 1993
;      APPLICATION NUMBER:  U.S. 08/009,389
;      FILING DATE:  23 February, 1993
;      APPLICATION NUMBER:  U.S. 08/017,127
;      FILING DATE:  12 February, 1993
;      APPLICATION NUMBER:  U.S. 07/934,161
;      FILING DATE:  21 August, 1992
;      APPLICATION NUMBER:  U.S. 07/834,044
;      FILING DATE:  11 February, 1992
;      APPLICATION NUMBER:  U.S. 07/749,451
;      FILING DATE:  23 August, 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Heber, Sheldon O.
;      REGISTRATION NUMBER:  38,179
;      REFERENCE/DOCKET NUMBER:  214/101
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1078 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-484-159-7

  Query Match             23.6%;  Score 1068.5;  DB 2;  Length 1078;
  Best Local Similarity   30.3%;  Pred. No. 2.5e-93;
  Matches  269;  Conservative  156;  Mismatches  366;  Indels   97;  Gaps   26;

Qy         12 WALL----HPGTGAPLCLSQQLRMKGDYVLGGLFPLG---EAEEAGLRSRTRPSSPVCTR 64
              | ||    |      |    |: : ||| :||||||:     |::  |:|  || |  | |
Db          8 WVLLALTWHTSAYGP---DQRAQKKGDIILGGLFPIHFGVAAKDQDLKS--RPESVECIR 62

Qy         65 FSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLA--KAGS 122
              ::  |  |  ||  |:||||:   ||| | ||| :||||:    |:: :| |:|  | |
Db         63 YNFRGFRWLQAMIFAIEEINSSPALLPNLTLGYRIFDTCNTVSKALEATLSFVAQNKIDS 122

Qy        123 RDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETFPS 182
               ::  :|| ::: |  :||:|   | ::          | :||||| :|  ||| :  | |
Db        123 LNLDEFCNCSEHIPSTIAVVGATGSGVSTAVANLLGLFYIPQVSYASSSRLLSNKNQFKS 182

Qy        183 FFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHEGL 242
              | ||:|:|  | || |:::: | ||||  : :||:||| |:  |   |  | |||    |
Db        183 FLRTIPNDEHQATAMADIIEYFRWNWVGTIAADDDYGRPGIEKFREEAEERDICIDFSEL 242

Qy        243 VPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVASEA 302
              :   :  |   ::| |:  :  |: :|:::|:|       |    :  :: |:|:|||
Db        243 I----SQYSDEEEIQHVVEVIQNSTAKVIVVFSSGPDLEPLIKEIVRRNITGKIWLASEA 298

Qy        303 WLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYV-KTHLALATDPAFCSALGERE--- 358
              | :| |:        :|  :||  :  |:  | ::: | |   :    |     |
Db        299 WASSSLIAMPQYFHVVGGTIGFALKAGQIPGFREFLKKVHPRKSVHNGFAKEFWEETFNC 358

Qy        359 ---------------QGLEEDVVGQRCPQCD------CITLQNVSA------GLNHHQ- 389
                           :| ||   | |           |   :|:|:         | :
Db        359 HLQEGAKGPLPVDTFLRGHEES--GDRFSNSSTAFRPLCTGDENISSVETPYIDYTHLRI 416

Qy        390 TFSVYAAVYSVAQALHNTLQC--------NASGCPAQDPVKPWQLLENMYNLTF-HVGGL 440
              :::|| ||||:| || :   |          | | |      |: ||:|::: :| | :  |
Db        417 SYNVYLAVYSIAHALQDIYTCLPGRGLFTNGS-CADIKKVEAWQVLKHLRHLNFTNNMGE 475

Qy        441 PLRFDSSGNVDMEYDLKLW---VWQGSVPRLHDVGRFN-GSLRTERL-----KIRWHTSD 491
              : ||  |::   | :  |     ||:   :|| :|  : : |||     || |
Db        476 QVTFDECGDLVGNYSIINWHLSPEDGSIV-FKEVGYYNVYAKKGERLFINEEKILWSGFS 534

Qy        492 NQKPVSRCSRQCQEGQVRR-VKGFHSCCYDCVDCEAGSYRQNPDDIACTFCGQDEWSPER 550
              : | | ||| |  |  :  ::|  :||::||:|  | |    |  ||  |  | || |
Db        535 REVPFSNCSRDCLAGTRKGIIEGEPTCCFECVECPDGEYSDETDASACNKCPDDFWSNEN 594

Qy        551 STRCFRRRSRFLAWGEPAVLLLLLLSLALGLVLAALGLFVHHRDSPLVQASGGPLA--- 607
              | |  :   ||:| ||  : | |   | : |    ||:|:  |::|:|:|:   |:
Db        595 HTSCIAKEIEFLSWTEPFGIALTLFAVLGIFLTAFVLGVFIKFRNTPIVKATNRELSYLL 654

Qy        608 CFGLVCLGLVCL-SVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV--ESEL 664
              | |:|    |  | | | |:|    |  :||   :    |:| : ::   : :  |:::
Db        655 LFSLLC----CFSSSLFFIGEPQDWTCRLRQPAFGISFVLCISCILVKTNRVLLVFEAKI 710

Qy        665 PLS-----WADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEAL 719
              | |     |  |          :|:| |   ::: :|  :|   ||     :  :
Db        711 PTSFHRKWWGLNLQ--------FLLVFLCTFMQIVICVIWLYTAPPSSYRNQELEDEIIF 762

Qy        720 VHCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFV 779
              : |   | :: |          || :||   |  |  |  :| |: :||:|| :|| |:||:
Db        763 ITCHEGSLMALGFLIGYTCLLAAICFFFAFKSRKLPENFNEAKFITFSMLIFFIVWISFI 822

Qy        780 PLLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPE 827
              |  |:       ||:: |:|    |:||     : |::: :|  || |
Db        823 PAYASTYGKFVSAVEVIAILAASFGLLACIFFNKIYIILFKPSRNTIE 870


RESULT 11
US-10-125-772-28
; Sequence 28, Application US/10125772
; Patent No. 6951739
; GENERAL INFORMATION:
;  APPLICANT: MariCal
;  APPLICANT:  Harris, H. William
;  APPLICANT:  Nearing, Jacqueline A.
;  APPLICANT:  Betka, Marlies
;  TITLE OF INVENTION: Polyvalent Cation-Sensing Receptor in Atlantic Salmon
;  FILE REFERENCE: 2213.1006-006
```

```
;  CURRENT APPLICATION NUMBER: US/10/125,772
;  CURRENT FILING DATE:  2002-08-16
;  PRIOR APPLICATION NUMBER: 10/121,441
;  PRIOR FILING DATE: 2002-04-11
;  PRIOR APPLICATION NUMBER: PCT/US01/31704
;  PRIOR FILING DATE: 2001-10-11
;  PRIOR APPLICATION NUMBER: 60/240,392
;  PRIOR FILING DATE: 2000-10-12
;  PRIOR APPLICATION NUMBER: 60/240,003
;  PRIOR FILING DATE: 2000-10-12
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
;   LENGTH: 1078
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-125-772-28

  Query Match           23.6%;  Score 1067.5;  DB 2;  Length 1078;
  Best Local Similarity   30.4%;  Pred. No. 3.1e-93;
  Matches  270;  Conservative  153;  Mismatches  368;  Indels   97;  Gaps   25;

Qy         12 WALL----HPGTGAPLCLSQQLRMKGDYVLGGLFPLG---EAEEAGLRSRTRPSSPVCTR 64
              | ||    |    :|   ||  : ||| :||||||:     |::  |:|  || |  | |
Db          8 WVLLALTWHTSAYSP---SQPAQKKGDIILGGLFPIHFGVAAKDQDLKS--RPESVECIR 62

Qy         65 FSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLA--KAGS 122
              ::  |  |  ||  |:||||:   ||| | ||| :||||:    |:: :| |:|  |  |
Db         63 YNFRGFRWLQAMIFAIEEINSSPALLPNLTLGYRIFDTCNTVSKALEATLSFVAQNKIDS 122

Qy        123 RDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETFPS 182
               ::  :|| ::: |  :||:|   | ::          | :||||| :|  ||| :  | |
Db        123 LNLDEFCNCSEHIPSTIAVVGATGSGVSTAVANLLGLFYIPQVSYASSSRLLSNKNQFKS 182

Qy        183 FFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHEGL 242
              | ||:|:|     || |:::: | ||||  : :||:||| |:  |   |  | |||    |
Db        183 FLRTIPNDEHCATAMADIIEYFRWNWVGTIAADDDYGRPGIEKFREEAEERDICIDFSEL 242

Qy        243 VPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVASEA 302
              :    : |   ::| |:  :  |: :|:::|:|       |    :   :: |:|:|||| 
Db        243 I----SQYSDEEEIQMVVEVIQNSTAKVIVVFSSGPDLEPLIKEIVPRNITGKIWLASEA 298

Qy        303 WLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYV-KTHLALATDPAFCSALGERE--- 358
              | :| |:        :|  :|| :  |:  | ::: | |   : :  |      
Db        299 WASSSLIAMPQYFHVVGGTIGFALKAGQIPGFREFLKKVHPPKSVNNGFAKEFWEETFMC 358

Qy        359 ----------------QGLEEDVVGQRCPQCD------CITLQNVSA-------GLNHHQ 389
                           | ||   | |          |   :|:|:          |
Db        359 HLQEGAKGPLPVDTFLAGHEES--GDRFSNSSTAFPPLCTGDENISSVETPYIDYTNLRI 416

Qy        390 TFSVYAAVYSVAQALHNTLQC--------NASGCPAQDPVKPWQLLENMYNLTF-HVGGL 440
              :::|| ||||:| || :   |         | | |      |: ||:|::: || | :  |
Db        417 SYNVYLAVYSIANALQDIYTCLPGRGLFTNGS-CADIKKVEAWQVLKHLRNLNFTNNMGE 475

Qy        441 PLRFDSSGNVDMEYDLKLW---VWQGSVPRLHDVGRFN-GSLRTERL-----KIRWHTSD 491
              : ||  |::   | :  |       ||:    :|| :|  : : |||     || |
Db        476 QVTFDECGDLVGNYSIINWHLSPEDGSIV-FKEVGYYNVYAKKGERLFINEEKILWSGFS 534

Qy        492 NQKPVSRCSRQCQEGQVRR-VKGFHSCCYDCVDCEAGSYRQNPDDIACTFCGQDEWSPER 550
              : | | |||  |  |  :  ::|  :||::||:|  | |    |  ||  |  | ||| |
Db        535 REVPFSNCSRDCLAGTRKGIIEGEPTCCFECVECPDGEYSDETDASACNKCPDDFWSNEN 594

Qy        551 STRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQASGGPLA--- 607
              | |  :   ||:| ||  : | |    | : |    ||:|:  |::|:|:|:   |:
Db        595 HTSCIAKEIEFLSWTEPFGIALTLFAVLGISLTAFVLGVFIKFRNTPIVKATNRELSYLL 654

Qy        608 CFGLVCLGLVCL-SVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV--ESEL 664
              | |:|    |  | | | |:|    |  :||   :   |:| : ::   :  :  |:::
Db        655 LFSLLC----CFSSSLFFIGEPQDWTCRLRQPAFGISFVLCISCILVKTNRVLLVFEAKI 710

Qy        665 PLS-----WADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEAL 719
              | |     |   |         :|:| |    :: :|  :|  ||    :  :
Db        711 PTSFMFKWWGLNLQ--------FLLVFLCTFNQIVICVIWLYTAPPSSYRNQELEDEIIF 762

Qy        720 VHCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFV 779
              : |   | :: |         || :||   |  |  |  :| |: :||:|| :|| |:||:
Db        763 ITCNEGSLMALGFLIGYTCLLAAICFFFAFKSRKLPENFNEAKPITFSMLIFFIVWISFI 822

Qy        780 PLLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPE 827
              |  |:        ||:: |:|    :|||     : |::: :|  || |
Db        823 PAYASTYGKFVSAVEVIAILAASFGLLACIFFNKIYIILFKPSRNTIE 870


RESULT 12
US-10-125-772-8
; Sequence 8, Application US/10125772
; Patent No. 6951739
; GENERAL INFORMATION:
;  APPLICANT: MariCal
;  APPLICANT:  Harris, H. William
;  APPLICANT:  Nearing, Jacqueline A.
;  APPLICANT:  Betka, Marlies
;  TITLE OF INVENTION: Polyvalent Cation-Sensing Receptor in Atlantic Salmon
;  FILE REFERENCE: 2213.1006-006
;  CURRENT APPLICATION NUMBER: US/10/125,772
;  CURRENT FILING DATE:  2002-08-16
;  PRIOR APPLICATION NUMBER: 10/121,441
;  PRIOR FILING DATE: 2002-04-11
;  PRIOR APPLICATION NUMBER: PCT/US01/31704
;  PRIOR FILING DATE: 2001-10-11
;  PRIOR APPLICATION NUMBER: 60/240,392
;  PRIOR FILING DATE: 2000-10-12
;  PRIOR APPLICATION NUMBER: 60/240,003
;  PRIOR FILING DATE: 2000-10-12
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
;   LENGTH: 941
;   TYPE: PRT
;   ORGANISM: Salmo salar
US-10-125-772-8

  Query Match           23.5%;  Score 1065;  DB 2;  Length 941;
  Best Local Similarity   30.4%;  Pred. No. 4.4e-93;
  Matches  264;  Conservative  158;  Mismatches  343;  Indels  102;  Gaps   26;

Qy         27 QQLRMKGDYVLGGLFPLGEAEEAGLRSR-----TRPSSPVCTRFSSNGLLWALAMKMAVE 81
              |: :  || :||||||:       |: |:      || |  | |::  |  |  ||  |:|
Db         24 QRAQKTGDILLGGLFPM----HFGVTSKDQDLAARPESTECVRYNFRGFRWLQAMIFAIE 79

Qy         82 EINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLA--KAGSRDIAAYCNYTQYQPRVL 139
              |||| | ||| : ||| :||||:    |:: :| |:|  |  | ::  :|| | : |  :
Db         80 EINNSSTLLPNITLGYRIFDTCNTVSKALEATLSFVAQNKIDSLNLDEFCNCTDHIPSTI 139

Qy        140 AVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETFPSFFRTVPSDRVQLTAAAE 199
              ||:|    | ::          | :|||:|| :|  |||  :  | || ||:|:|  | || |:
Db        140 AVVGASGSAVSTAVANLLGLFYIPQISYASSSRLLSNKNQFKSFMRTIPTDEHQATAMAD 199

Qy        200 LLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHEGLVPLPRADDSRLGKVQDV 259
              ::  | |||| |: |||||||  |:  |       | |||     |:    :       ::| :
Db        200 IIDYFQWNWVIAVASDDEYGRPGIEKFEKEMEERDICIHLSELI----SQYFEEWQIQGL 255

Qy        260 LHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVASEAWLTS---------DLVM 310
              : ::  || :|:::|||       |    :   :: ::|:|||||  |:        |:|:
Db        256 VDRIENSSAKVIVVFASGPDIEPLIKEMVRRNITDRIWLASEAWATTSLIAKPEYLDVVV 315

Qy        311 GLPGMA-QMGTVLG---FLQ----RGAQLHEFPQ--YVKTHLALATDPAFCSALGEREQG 360
              |  | | : | :  |   |||    :  :  :|| :  : :|       |       | : | |
Db        316 GTIGFALRAGEIPGFKDFLQEVTPKKSSHNEFVREFWEETFNCYLED---SQRLRDSENG 372

Qy        361 L---------EEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQC- 410
                        |||::|   |  |   |:          :::|| ||:|:|||| : | |
Db        373 STSFRPLCTGEEDIMGAETPYLDYTHLR---------ISYNVYVAVHSIAQALQDILTCI 423

Qy        411 ------NASGCPAQDPVKPWQLLENMYNLTF-HVGGLPLRFDSSGNVDMEYDLKLW---V 460
                   : : |      :: ||:|: : :| | :  |  : || : :    | :  |
```

```
Db        424 PGRGLFSNNSCADIKKIEAWQVLKQLRHLNFSNSMGEKVHFDENADPSGNYTIINWHRSP 483

Qy        461 WQGSVPRLHDVGRFNGSLRTER--------LKIRWHTSDNQKPVSRCSRQCQEGQVRR-V 511
              |||    :|| :|  :| :|          || |:  : : | | ||  |: |  :  :
Db        484 EDGSVV-FEEVGFYN--MRAKRGVQLFIDNTKILWNGYNTEVPFSNCSEDCEPGTRKGII 540

Qy        512 KGFHSCCYDCVDCEAGSYRQNPDDIACTFCGQDEWSPERSTRCFRRRSRFLAWGEPAVLL 571
              :   :||::| :|  | |  : |   || |  | || |  | || :   ||:| ||  :
Db        541 ESMPTCCFECTECSEGEYSDHKDASVCTKCPNDSWSNENHTSCFLKEIEFLSWTEPFGIA 600

Qy        572 LLLLLSLALGLVLAALGLFVHHRDSPLVQASGGPLA---CFGLVCLGLVCL-SVLLFPGQ 627
              | |   | : |    :|:|:  |::|:|:|:   |:    | |:|    |  | |:| |:
Db        601 LALCSVLGVFLTAFVMGVFIKFRNTPIVKATNRELSYLLLFSLIC----CFSSSLIFIGE 656

Qy        628 PSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV--ESELPLSWADRLSGCLRGPW----- 680
              |    |  :||   :    |:| : ::   : :  |:::| |          |   |
Db        657 PQDWTCRLRQPAFGISFVLCISCILVKTNRVLLVFEAKIPTS--------LHRKWWGLNL 708

Qy        681 AWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALVHCRTRSWVSFGLAHATNATL 740
              :|:| |   |:| :|  :|   ||    : | :      :  |   | :: |       |
Db        709 QFLLVFLFTFVQVMICVVWLYNAPPASYRN-HDIDEIIFITCNEGSMMALGFLIGYTCLL 767

Qy        741 AFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVPLLANVQVVLRPAVQMGALLL 800
              | :||   |  |  |  :  |: :||:|| :|| |:||:|   :        ||:: |:|
Db        768 AAICFFFAFKSRKLPENFTEAKFITFSMLIFFIVWISFIPAYFSTYGKFVSAVEVIAILA 827

Qy        801 CVLGILAAFHLPRCYLLMRQPGLNTPE 827
                |:||     : |::: :|  || |
Db        828 SSFGLLACIFFNKVYIILFKPSRNTIE 854


RESULT 13
US-10-125-772-10
; Sequence 10, Application US/10125772
; Patent No. 6951739
; GENERAL INFORMATION:
;  APPLICANT: MariCal
;  APPLICANT:  Harris, H. William
;  APPLICANT:  Nearing, Jacqueline A.
;  APPLICANT:  Betka, Marlies
;  TITLE OF INVENTION: Polyvalent Cation-Sensing Receptor in Atlantic Salmon
;  FILE REFERENCE: 2213.1006-006
;  CURRENT APPLICATION NUMBER: US/10/125,772
;  CURRENT FILING DATE:  2002-08-16
;  PRIOR APPLICATION NUMBER: 10/121,441
;  PRIOR FILING DATE: 2002-04-11
;  PRIOR APPLICATION NUMBER: PCT/US01/31704
;  PRIOR FILING DATE: 2001-10-11
;  PRIOR APPLICATION NUMBER: 60/240,392
;  PRIOR FILING DATE: 2000-10-12
;  PRIOR APPLICATION NUMBER: 60/240,003
;  PRIOR FILING DATE: 2000-10-12
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
;   LENGTH: 941
;   TYPE: PRT
;   ORGANISM: Salmo salar
US-10-125-772-10

  Query Match           23.5%;  Score 1064;  DB 2;  Length 941;
  Best Local Similarity   30.4%;  Pred. No. 5.4e-93;
  Matches  264;  Conservative  158;  Mismatches  343;  Indels  102;  Gaps   26;

Qy         27 QQLRMKGDYVLGGLFPLGEAEEAGLRSR-----TRPSSPVCTRFSSNGLLWALAMKMAVE 81
              |: :  || :||||||:      |: |:       || |  | |::  |  |  ||  |:|
Db         24 QRAQKTGDILLGGLFPM----HFGVTSKDQDLAARPESTECVRYNFRGFRWLQAMIFAIE 79

Qy         82 EINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLA--KAGSRDIAAYCNYTQYQPRVL 139
              |||| | ||| : ||| :||||:    |:: :| |:|  |  | ::  :|| | : |  :
Db         80 EINNSSTLLPNITLGYRIFDTCNTVSKALEATLSFVAQNKIDSLNLDEFCNCTDHIPSTI 139

Qy        140 AVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETFPSFFRTVPSDRVQLTAAAE 199
              ||:|   | ::         | :||:|| :|  ||| :  | || ||:|:|  | || |:
Db        140 AVVGASGSAVSTAVANLLGLFYIPQISYASSSRLLSNKNQFKSFMRTIPTDEHQATAMAD 199

Qy        200 LLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHEGLVPLPRADDSRLGKVQDV 259
              ::  | |||| |: |||||| |:  |      | |||    |:    :      ::| :
Db        200 IIDYFQWNWVIAVASDDEYGRPGIEKFEKEMEERDICIHLSELI----SQYFEEWQIQGL 255

Qy        260 LHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVASEAWLTS---------DLVM 310
              : ::  || :|:::|||      |    :   :: ::|:|||| |:          |:|:
Db        256 VGRIENSSAKVIVVFASGPDIEPLIKEMVRRNITDRIWLASEAWATTSLIAKPEYLDVVV 315

Qy        311 GLPGMA-QMGTVLG---FLQ----RGAQLHEFPQ--YVKTHLALATDPAFCSALGEREQG 360
              |  | | : | : |   |||    : :  :|| :  : :|       |       | : | |
Db        316 GTIGFALRAGEIPGFKDFLQEVTPKKSSHNEFVREFWEETFNCYLED---SQRLRDSENG 372

Qy        361 L---------EEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQC- 410
                      |||::|   |  |   |:          :::|| ||:|:|||| : | |
Db        373 STSFRPLCTGEEDIMGAETPYLDYTHLR---------ISYNVYVAVHSIAQALQDILTCI 423

Qy        411 ------NASGCPAQDPVKPWQLLENMYNLTF-HVGGLPLRFDSSGNVDMEYDLKLW---V 460
                   : : |      :: ||:|: : :| | :  |  : || : :    | :  |
Db        424 PGRGLFSNNSCADIKKIEAWQVLKQLRHLNFSNSMGEKVHFDENADPSGNYTIINWHRSP 483

Qy        461 WQGSVPRLHDVGRFNGSLRTER--------LKIRWHTSDNQKPVSRCSRQCQEGQVRR-V 511
              |||    :|| :|  :| :|          || |:  : : | | ||  |: |  :  :
Db        484 EDGSVV-FEEVGFYN--MRAKRGVQLFIDNTKILWNGYNTEVPFSNCSEDCEPGTRKGII 540

Qy        512 KGFHSCCYDCVDCEAGSYRQNPDDIACTFCGQDEWSPERSTRCFRRRSRFLAWGEPAVLL 571
              :   :||::| :|  | |  : |   || |  | || |  | || :   ||:| ||  :
Db        541 ESMPTCCFECTECSEGEYSDHKDASVCTKCPNDSWSNENHTSCFLKEIEFLSWTEPFGIA 600

Qy        572 LLLLLSLALGLVLAALGLFVHHRDSPLVQASGGPLA---CFGLVCLGLVCL-SVLLFPGQ 627
              | |   | : |    :|:|:  |::|:|:|:   |:    | |:|    |  | |:| |:
Db        601 LALCSVLGVFLTAFVMGVFIKFRNTPIVKATNRELSYLLLFSLIC----CFSSSLIFIGE 656

Qy        628 PSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV--ESELPLSWADRLSGCLRGPW----- 680
              |    |  :||   :    |:| : ::   : :  |:::| |          |   |
Db        657 PQDWTCRLRQPAFGISFVLCISCILVKTNRVLLVFEAKIPTS--------LHRKWWGLNL 708

Qy        681 AWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALVHCRTRSWVSFGLAHATNATL 740
              :|:| |   |:| :|  :|   ||    : | :      :  |   | :: |       |
Db        709 QFLLVFLFTFVQVMICVVWLYNAPPASYRN-HDIDEIIFITCNEGSMMALGFLIGYTCLL 767

Qy        741 AFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVPLLANVQVVLRPAVQMGALLL 800
              | :||   |  |  |  :  |: :||:|| :|| |:||:|   :        ||:: |:|
Db        768 AAICFFFAFKSRKLPENFTEAKFITFSMLIFFIVWISFIPAYFSTYGKFVSAVEVIAILA 827

Qy        801 CVLGILAAFHLPRCYLLMRQPGLNTPE 827
                |:||     : |::: :|  || |
Db        828 SSFGLLACIFFNKVYIILFKPSRNTIE 854


RESULT 14
US-08-485-588-5
; Sequence 5, Application US/08485588
; Patent No. 5688938
;  GENERAL INFORMATION:
;    APPLICANT:  Edward M. Brown
;    APPLICANT:  Steven C. Hebert
;    APPLICANT:  Forrest H. Fuller
;    APPLICANT:  James E. Garrett, Jr.
;    TITLE OF INVENTION:  CALCIUM RECEPTOR-ACTIVE
;    TITLE OF INVENTION:  MOLECULES
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  First Interstate World Center
;      STREET:  Suite 4700
;      STREET:  633 West Fifth Street
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90071
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb storage
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FASTSEQ
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/485,588
;      FILING DATE:  7 June, 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;    PRIOR APPLICATION DATA:  including application
;    PRIOR APPLICATION DATA:  described below:  9
;      APPLICATION NUMBER:  08/353,784
;      FILING DATE:  9 December, 1994
;      APPLICATION NUMBER:  PCT/US/94/12117
;      FILING DATE:  21 October, 1994
;      APPLICATION NUMBER:  U.S. 08/292,827
;      FILING DATE:  23 August, 1994
;      APPLICATION NUMBER:  U.S. 08/141,248
;      FILING DATE:  22 October, 1993
;      APPLICATION NUMBER:  U.S. 08/009,389
;      FILING DATE:  23 February, 1993
;      APPLICATION NUMBER:  U.S. 08/017,127
;      FILING DATE:  12 February, 1993
;      APPLICATION NUMBER:  U.S. 07/934,161
;      FILING DATE:  21 August, 1992
;      APPLICATION NUMBER:  U.S. 07/834,044
;      FILING DATE:  11 February, 1992
;      APPLICATION NUMBER:  U.S. 07/749,451
;      FILING DATE:  23 August, 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Heber, Sheldon O.
;      REGISTRATION NUMBER:  38,179
;      REFERENCE/DOCKET NUMBER:  213/005
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1085 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-485-588-5

  Query Match             23.4%;  Score 1060;  DB 1;  Length 1085;
  Best Local Similarity   30.0%;  Pred. No. 1.7e-92;
  Matches  269;  Conservative  155;  Mismatches  362;  Indels  112;  Gaps   26;

Qy          6 VLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPL--GEA-EEAGLRSRTRPSSPVC 62
              :|  | |      |        |: : ||| :|||||:  | | ::  |:|  || |  |
Db         10 LLAFSTWCTSAYGP------DQRAQKKGDIILGGLFPIHFGVAVKDQDLKS--RPESVEC 61

Qy         63 TRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLA--KA 120
              |::  |  |  ||  |:||||:   ||| : ||| :||||:    |:: :| |:|  |
Db         62 IRYNFRGFRWLQAMIFAIEEINSSPALLPNMTLGYRIFDTCNTVSKALEATLSFVAQNKI 121

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              | ::  :|| ::: |  :||:|   | ::          | :||||| :|  ||| :  |
Db        122 DSLNLDEFCNCSEHIPSTIAVVGATGSGISTAVANLLGLFYIPQVSYASSSRLLSNKNQF 181

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              || ||:|:|  | || |:::: | |||| : :||:||| |:  |    |  | |||
Db        182 KSFLRTIPNDEHQATAMADIIEYFRWNWVGTIAADDDYGRPGIEKFREEAEERDICIDFS 241

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              |:    :  |   |:| |:  :  |: :|:::|:|      |    :   :: ::|:||
Db        242 ELI----SQYSDEEKIQQVVEVIQNSTAKVIVVFSSGPDLEPLIKEIVRRNITGRIWLAS 297

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYV-KTHLALATDPAFCSAL----- 354
              ||| :| |:        :|  :||  :  |:  | ::: | |   :    |
Db        298 EAWASSSLIAMPEYFHVVGGTIGFGLKAGQIPGFREFLQKVHPRKSVHNGFAKEFWEETF 357

Qy        355 ------------------GEREQGL--------------EEDVVGQRCPQCDCITLQN 380
                               |  | ||                ||::     |  |   |:
Db        358 NCHLQEGAKGPLPVDTFLRGHEEGGARLSNSPTAFRPLCTGEENISSVETPYMDYTHLR- 416

Qy        381 VSAGLNHHQTFSVYAAVYSVAQALHNTLQC--------NASGCPAQDPVKPWQLLENMYN 432
                    :::|| ||||:| || :  |         | | |     |: ||:|::: :
Db        417 --------ISYNVYLAVYSIAHALQDIYTCIPGRGLFTNGS-CADIKKVEAWQVLKHLRH 467

Qy        433 LTFHVG-GLPLRFDSSGNVDMEYDLKLW---VWQGSVPRLHDVGRFN-GSLRTERL---- 483
              | |    |  : ||  |::   | :  |       ||:    :|| :|  : : |||
Db        468 LNFTSNMGEQVTFDECGDLAGNYSIINWHLSPEDGSIV-FKEVGYYNVYAKKGERLFIND 526

Qy        484 -KIRWHTSDNQKPVSRCSRQCQEGQVRR-VKGFHSCCYDCVDCEAGSYRQNPDDIACTFC 541
               || |     : | | ||| |  |  :  ::|  :||::||:|  | |    |  ||  |
Db        527 EKILWSGFSREVPFSNCSRDCLAGTRKGIIEGEPTCCFECVECPDGEYSDETDASACDKC 586

Qy        542 GQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQA 601
              | || |  | |  :   ||:| ||  : | |   | : |     ||:|:  |::|:|:|
Db        587 PDDFWSNENHTSCIAKEIEFLSWTEPFGIALTLFAVLGIFLTAFVLGVFIKFRNTPIVKA 646

Qy        602 SGGPLA---CFGLVCLGLVCL-SVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAE 657
              :   |:    | |:|    |  | | | |:|    |  :||   :    |:| : ::
Db        647 TNRELSYLLLFSLLC----CFSSSLFFIGEPQDWTCRLRQPAFGISFVLCISCILVKTNR 702

Qy        658 IFV--ESELPLS-----WADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTD 710
              : :  |:::| |     |  |          :|:| |   ::: :|  :|   ||    :
Db        703 VLLVFEAKIPTSFHRKWWGLNLQ--------FLLVFLCTFMQIVICAIWLNTAPPSSYRN 754

Qy        711 WHMLPTEAL-VHCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAML 769
              | |  | : : |   | :: |         || :||   |  |  |  :| |: :||:||
Db        755 -HELEDEIIFITCHEGSLMALGFLIGYTCLLAAICFFFAFKSRKLPENFNEAKFITFSML 813

Qy        770 AYFITWVSFVPLLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPE 827
              :|| |:||:|  |:       ||:: |:|    |:||     : |::: :|  || |
Db        814 IFFIVWISFIPAYASTYGKFVSAVEVIAILAASFGLLACIFFNKVYIILFKPSRNTIE 871


RESULT 15
US-08-484-565-5
; Sequence 5, Application US/08484565
; Patent No. 5763569
;  GENERAL INFORMATION:
;    APPLICANT:  Edward M. Brown
;    APPLICANT:  Steven C. Hebert
;    APPLICANT:  James E. Garrett, Jr.
;    TITLE OF INVENTION:  CALCIUM RECEPTOR-ACTIVE
;    TITLE OF INVENTION:  MOLECULES
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  First Interstate World Center
;      STREET:  Suite 4700
;      STREET:  633 West Fifth Street
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb storage
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FASTSEQ
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,565
;      FILING DATE:  7 June, 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;    PRIOR APPLICATION DATA:  including application
;    PRIOR APPLICATION DATA:  described below:  9
;      APPLICATION NUMBER:  08/353,784
;      FILING DATE:  9 December, 1994
;      APPLICATION NUMBER:  PCT/US/94/12117
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2005, 17:56:20 ; Search time 236 Seconds
(without alignments)
2547.077 Million cell updates/sec

Title: US-10-679-102-15
Perfect score: 4524
Sequence: 1 MLGPAVLGLSLWALLHPGTG.....GPGDAQGQNDGNTGNQGKHE 852

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 4524 | 100.0 | 852 | 1 | TS1R3_HUMAN | Q7rtx0 homo sapien |
| 2 | 4472 | 98.9 | 852 | 1 | TS1R3_PANTR | Q717c2 pan troglod |
| 3 | 4439 | 98.1 | 852 | 1 | TS1R3_GORGO | Q717c1 gorilla gor |
| 4 | 3255 | 71.9 | 858 | 1 | TS1R3_RAT | Q923k1 rattus norv |
| 5 | 3229 | 71.4 | 858 | 1 | TS1R3_MOUSE | Q925d8 mus musculu |
| 6 | 1483.5 | 32.8 | 829 | 2 | Q4RX45_TETNG | Q4rx45 tetraodon n |
| 7 | 1232 | 27.2 | 857 | 2 | Q4RX46_TETNG | Q4rx46 tetraodon n |
| 8 | 1208.5 | 26.7 | 840 | 1 | TS1R1_RAT | Q9z0r8 rattus norv |
| 9 | 1180.5 | 26.1 | 842 | 1 | TS1R1_MOUSE | Q99pg6 mus musculu |
| 10 | 1153 | 25.5 | 841 | 1 | TS1R1_HUMAN | Q7rtx1 homo sapien |
| 11 | 1095.5 | 24.2 | 928 | 2 | Q8K4Z6_MOUSE | Q8k4z6 mus musculu |
| 12 | 1086.5 | 24.0 | 928 | 2 | Q70VB1_RAT | Q70vb1 rattus norv |
| 13 | 1072 | 23.7 | 940 | 2 | Q5YEV6_OREMO | Q5yev6 oreochromis |
| 14 | 1068.5 | 23.6 | 1078 | 1 | CASR_HUMAN | P41180 homo sapien |
| 15 | 1068.5 | 23.6 | 1078 | 2 | Q4PJ19_HUMAN | Q4pj19 homo sapien |
| 16 | 1065 | 23.5 | 941 | 2 | Q6XAF1_SALSA | Q6xaf1 salmo salar |
| 17 | 1064 | 23.5 | 941 | 2 | Q6XAF3_SALSA | Q6xaf3 salmo salar |
| 18 | 1062 | 23.5 | 1079 | 1 | CASR_MOUSE | Q9qy96 mus musculu |
| 19 | 1061.5 | 23.5 | 926 | 2 | Q5T6X5_HUMAN | Q5t6x5 homo sapien |
| 20 | 1060 | 23.4 | 940 | 2 | Q90WL6_SPAAU | Q90wl6 sparus aura |
| 21 | 1060 | 23.4 | 1085 | 1 | CASR_BOVIN | P35384 bos taurus |
| 22 | 1056 | 23.3 | 1027 | 2 | Q8JI04_SQUAC | Q8ji04 squalus aca |
| 23 | 1054 | 23.3 | 1079 | 1 | CASR_RAT | P48442 rattus norv |
| 24 | 1053.5 | 23.3 | 926 | 2 | Q8NHZ9_HUMAN | Q8nhz9 homo sapien |
| 25 | 1051 | 23.2 | 1079 | 2 | Q80ZA8_RAT | Q80za8 rattus norv |
| 26 | 1049 | 23.2 | 940 | 2 | O73635_FUGRU | O73635 fugu rubrip |
| 27 | 1048.5 | 23.2 | 850 | 2 | Q6XAF2_SALSA | Q6xaf2 salmo salar |
| 28 | 1021.5 | 22.6 | 879 | 2 | Q8NGU8_HUMAN | Q8ngu8 homo sapien |
| 29 | 1017 | 22.5 | 841 | 2 | Q4S813_TETNG | Q4s813 tetraodon n |
| 30 | 1014.5 | 22.4 | 839 | 1 | TS1R2_HUMAN | Q8te23 homo sapien |
| 31 | 1014 | 22.4 | 857 | 2 | Q5DK51_MOUSE | Q5dk51 mus musculu |
| 32 | 1009.5 | 22.3 | 877 | 2 | Q9PW88_CARAU | Q9pw88 carassius a |
| 33 | 1002.5 | 22.2 | 843 | 1 | TS1R2_RAT | Q9z0r7 rattus norv |
| 34 | 1000.5 | 22.1 | 843 | 1 | TS1R2_MOUSE | Q925i4 mus musculu |
| 35 | 994 | 22.0 | 867 | 2 | Q5U9X3_BRARE | Q5u9x3 brachydanio |
| 36 | 977 | 21.6 | 855 | 2 | Q6JK43_HUMAN | Q6jk43 homo sapien |
| 37 | 973 | 21.5 | 818 | 2 | Q501X9_BRARE | Q501x9 brachydanio |
| 38 | 953 | 21.1 | 870 | 2 | Q4SFL1_TETNG | Q4sfl1 tetraodon n |
| 39 | 933.5 | 20.6 | 880 | 2 | O73639_FUGRU | O73639 fugu rubrip |
| 40 | 925 | 20.4 | 864 | 2 | O73637_FUGRU | O73637 fugu rubrip |
| 41 | 918.5 | 20.3 | 832 | 2 | Q5I5C9_9TELE | Q5i5c9 cyprinid sp |
| 42 | 910.5 | 20.1 | 838 | 2 | Q4S834_TETNG | Q4s834 tetraodon n |
| 43 | 899 | 19.9 | 753 | 2 | Q5DK52_MOUSE | Q5dk52 mus musculu |
| 44 | 892.5 | 19.7 | 828 | 2 | Q5I5D2_9TELE | Q5i5d2 cyprinid sp |
| 45 | 891 | 19.7 | 848 | 2 | Q5I5D0_9TELE | Q5i5d0 cyprinid sp |

ALIGNMENTS

RESULT 1
TS1R3_HUMAN
ID TS1R3_HUMAN STANDARD; PRT; 852 AA.
AC Q7RTX0; Q5TA49; Q8NGW9;
DT 01-FEB-2005 (Rel. 46, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Taste receptor type 1 member 3 precursor (Sweet taste receptor T1R3).
GN Name=TAS1R3; Synonyms=T1R3, TR3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Suwa M., Sato T., Okouchi I., Arita M., Futami K., Matsumoto S.,
RA Tsutsumi S., Aburatani H., Asai K., Akiyama Y.;
RT "Genome-wide discovery and analysis of human seven transmembrane helix
RT receptor genes.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RG Human chromosome 1 international sequencing consortium;
RL Submitted (DEC-2004) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP IDENTIFICATION, AND FUNCTION.
RX MEDLINE=21927605; PubMed=11917125; DOI=10.1073/pnas.072090199;
RA Li X., Staszewski L., Xu H., Durick K., Zoller M., Adler E.;
RT "Human receptors for sweet and umami taste.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4692-4696(2002).
RN [4]
RP FUNCTION.
RX PubMed=12892531; DOI=10.1290/1543-706X(2003)039<0080:TRTIAE>2.0.CO;2;
RA Ariyasu T., Matsumoto S., Kyono F., Hanaya T., Arai S., Ikeda M.,
RA Kurimoto M.;
RT "Taste receptor T1R3 is an essential molecule for the cellular
RT recognition of the disaccharide trehalose.";
RL In Vitro Cell. Dev. Biol. Anim. 39:80-88(2003).
RN [5]
RP REGION IMPLICATED IN SWEET RECEPTOR FUNCTION, AND MUTAGENESIS OF
RP ALA-537 AND PHE-540.
RX PubMed=15299024; DOI=10.1074/jbc.M406779200;
RA Jiang P., Ji Q., Liu Z., Snyder L.A., Benard L.M.J., Margolskee R.F.,
RA Max M.;
RT "The cysteine-rich region of T1R3 determines responses to intensely
RT sweet proteins.";
RL J. Biol. Chem. 279:45068-45075(2004).
CC -!- FUNCTION: Putative taste receptor. TAS1R1/TAS1R3 responds to the
CC umami taste stimulus (the taste of monosodium glutamate).
CC TAS1R2/TAS1R3 recognizes diverse natural and synthetic sweeteners.
CC TAS1R3 is essential for the recognition and response to the
CC disaccharide trehalose (By similarity). Sequence differences
CC within and between species can significantly influence the

```
CC       selectivity and specificity of taste responses.
CC   -!- SUBUNIT: Forms homodimers or heterodimers with TAS1R1 and TAS1R2.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to the G-protein coupled receptor 3 family.
CC       TAS1R subfamily.
CC   -!- DATABASE: NAME=Protein Spotlight; NOTE=Issue 55 of February 2005;
CC       WWW="http://www.expasy.org/spotlight/back_issues/sptlt055.shtml".
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; AB065647; BAC05873.1; -; Genomic_DNA.
DR   EMBL; AL139287; CAI23184.1; -; Genomic_DNA.
DR   EMBL; BK000152; DAA00013.1; -; mRNA.
DR   Ensembl; ENSG00000169962; Homo sapiens.
DR   HGNC; HGNC:15661; TAS1R3.
DR   MIM; 605865; -.
DR   GO; GO:0016021; C:integral to membrane; IC.
DR   GO; GO:0046982; F:protein heterodimerization activity; NAS.
DR   GO; GO:0008527; F:taste receptor activity; IDA.
DR   GO; GO:0050916; P:perception of sweet taste; IDA.
DR   GO; GO:0050917; P:perception of umami taste; IDA.
DR   InterPro; IPR001828; ANF_rcpt_lig_bd.
DR   InterPro; IPR000068; Ca_sens_receptor.
DR   InterPro; IPR000337; GPCR_Mgr.
DR   InterPro; IPR011500; NCD3G_GPCR.
DR   Pfam; PF00003; 7tm_3; 1.
DR   Pfam; PF01094; ANF_receptor; 1.
DR   Pfam; PF07562; NCD3G; 1.
DR   PRINTS; PR00592; CASENSINGR.
DR   PRINTS; PR00248; GPCRMGR.
DR   PROSITE; PS00979; G_PROTEIN_RECEP_F3_1; FALSE_NEG.
DR   PROSITE; PS00980; G_PROTEIN_RECEP_F3_2; 1.
DR   PROSITE; PS00981; G_PROTEIN_RECEP_F3_3; FALSE_NEG.
DR   PROSITE; PS50259; G_PROTEIN_RECEP_F3_4; 1.
KW   G-protein coupled receptor; Glycoprotein; Polymorphism; Receptor;
KW   Sensory transduction; Signal; Taste; Transducer; Transmembrane.
FT   SIGNAL       1     20       Potential.
FT   CHAIN       21    852       Taste receptor type 1 member 3.
FT   TOPO_DOM    21    570       Extracellular (Potential).
FT   TRANSMEM   571    591       1 (Potential).
FT   TOPO_DOM   592    603       Cytoplasmic (Potential).
FT   TRANSMEM   604    624       2 (Potential).
FT   TOPO_DOM   625    639       Extracellular (Potential).
FT   TRANSMEM   640    660       3 (Potential).
FT   TOPO_DOM   661    682       Cytoplasmic (Potential).
FT   TRANSMEM   683    703       4 (Potential).
FT   TOPO_DOM   704    729       Extracellular (Potential).
FT   TRANSMEM   730    750       5 (Potential).
FT   TOPO_DOM   751    762       Cytoplasmic (Potential).
FT   TRANSMEM   763    783       6 (Potential).
FT   TOPO_DOM   784    789       Extracellular (Potential).
FT   TRANSMEM   790    810       7 (Potential).
FT   TOPO_DOM   811    852       Cytoplasmic (Potential).
FT   REGION     536    545       Required for brazzein responsiveness.
FT   CARBOHYD    85     85       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   130    130       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   264    264       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   285    285       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   380    380       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   411    411       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   432    432       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   475    475       N-linked (GlcNAc. . .) (Potential).
FT   VARIANT    757    757       C -> R (in dbSNP:307377).
FT                               /FTId=VAR_020788.
FT   MUTAGEN    537    537       A->G: Retains partial activity toward
FT                               brazzein; however response to other
FT                               sweeteners tested is suppressed.
FT   MUTAGEN    537    537       A->P: Receptor unresponsive to all
FT                               sweeteners tested.
FT   MUTAGEN    537    537       A->T,S,E,V: Brazzein response abolished.
FT   MUTAGEN    540    540       F->A,H: Reduces the response to brazzein
FT                               and monellin.
FT   MUTAGEN    540    540       F->L: Reduces the response to monellin.
FT   MUTAGEN    540    540       F->Y,P: Reduces the response to brazzein;
FT                               P-540 also enhances responses to the
FT                               small molecule sweeteners.
FT   CONFLICT   852    852       E -> EAPILHPARHHSSGRPGGGTQAYAQGRARRGRGGTQ
FT                               LPTLPFAQWVDSIMTVTSTRDRAQDLTPARRQAHTGSPVRV
FT                               PGSAPSRA (in Ref. 1).
SQ   SEQUENCE   852 AA;  93386 MW;  9690DEC23E75EB3E CRC64;
```

```
  Query Match             100.0%;  Score 4524;  DB 1;  Length 852;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  852;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360

Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420

Qy        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540

Qy        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600

Qy        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720

Qy        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780

Qy        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840

Qy        841 NDGNTGNQGKHE 852
```

```
Db        841 NDGNTGNQGKHE 852


RESULT 2
TS1R3_PANTR
ID   TS1R3_PANTR    STANDARD;      PRT;   852 AA.
AC   Q717C2;
DT   01-FEB-2005 (Rel. 46, Created)
DT   01-FEB-2005 (Rel. 46, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Taste receptor type 1 member 3 precursor (Sweet taste receptor T1R3).
GN   Name=TAS1R3; Synonyms=T1R3;
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Li X., Reed D.R., Tordoff M.G., Bachmanov A.A., Mascioli K.J., Bak G.,
RA   Li W., Beauchamp G.K.;
RT   "Analysis of the T1R taste receptor genes in primates.";
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Putative taste receptor. TAS1R1/TAS1R3 responds to the
CC       umami taste stimulus (the taste of monosodium glutamate).
CC       TAS1R2/TAS1R3 recognizes diverse natural and synthetic sweeteners.
CC       TAS1R3 is essential for the recognition and response to the
CC       disaccharide trehalose (By similarity). Sequence differences
CC       within and between species can significantly influence the
CC       selectivity and specificity of taste responses (By similarity).
CC   -!- SUBUNIT: Forms homodimers or heterodimers with TAS1R1 and TAS1R2
CC       (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to the G-protein coupled receptor 3 family.
CC       TAS1R subfamily.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; AF545573; AAQ11896.1; -; Genomic_DNA.
DR   InterPro; IPR001828; ANF_rcpt_lig_bd.
DR   InterPro; IPR000068; Ca_sens_receptor.
DR   InterPro; IPR000337; GPCR_Mgr.
DR   InterPro; IPR011500; NCD3G_GPCR.
DR   Pfam; PF00003; 7tm_3; 1.
DR   Pfam; PF01094; ANF_receptor; 1.
DR   Pfam; PF07562; NCD3G; 1.
DR   PRINTS; PR00592; CASENSINGR.
DR   PRINTS; PR00248; GPCRMGR.
DR   PROSITE; PS00979; G_PROTEIN_RECEP_F3_1; FALSE_NEG.
DR   PROSITE; PS00980; G_PROTEIN_RECEP_F3_2; 1.
DR   PROSITE; PS00981; G_PROTEIN_RECEP_F3_3; FALSE_NEG.
DR   PROSITE; PS50259; G_PROTEIN_RECEP_F3_4; 1.
KW   G-protein coupled receptor; Glycoprotein; Receptor; Signal;
KW   Transducer; Transmembrane.
FT   SIGNAL        1     20       Potential.
FT   CHAIN        21    852       Taste receptor type 1 member 3.
FT   TOPO_DOM     21    570       Extracellular (Potential).
FT   TRANSMEM    571    591       1 (Potential).
FT   TOPO_DOM    592    603       Cytoplasmic (Potential).
FT   TRANSMEM    604    624       2 (Potential).
FT   TOPO_DOM    625    639       Extracellular (Potential).
FT   TRANSMEM    640    660       3 (Potential).
FT   TOPO_DOM    661    682       Cytoplasmic (Potential).
FT   TRANSMEM    683    703       4 (Potential).
FT   TOPO_DOM    704    729       Extracellular (Potential).
FT   TRANSMEM    730    750       5 (Potential).
FT   TOPO_DOM    751    762       Cytoplasmic (Potential).
FT   TRANSMEM    763    783       6 (Potential).
FT   TOPO_DOM    784    791       Extracellular (Potential).
FT   TRANSMEM    792    812       7 (Potential).
FT   TOPO_DOM    813    852       Cytoplasmic (Potential).
FT   CARBOHYD     85     85       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    130    130       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    264    264       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    285    285       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    380    380       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    411    411       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    432    432       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    475    475       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   852 AA;  93456 MW;  1E657EFDA949A98B CRC64;

  Query Match             98.9%;  Score 4472;  DB 1;  Length 852;
  Best Local Similarity   98.9%;  Pred. No. 2.8e-316;
  Matches  843;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLVFLAKA 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALAADPAFCSALGEREQG 360

Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420

Qy        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
              |||||||||||||| ||| ||||||||||||||||||||||||||||||||||||||||
Db        421 VKPWQLLENMYNLTFHAGGLMLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540

Qy        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFIHHRDSPLVQ 600

Qy        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720

Qy        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGRYNRARGLTFAMLAYFITWVSFVP 780

Qy        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||:
Db        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMWQPGLNTPEFFLGGGPGDAQGR 840
```

```
Qy        841 NDGNTGNQGKHE 852
              |||:||||||||
Db        841 NDGDTGNQGKHE 852


RESULT 3
TS1R3_GORGO
ID   TS1R3_GORGO    STANDARD;      PRT;   852 AA.
AC   Q717C1;
DT   01-FEB-2005 (Rel. 46, Created)
DT   01-FEB-2005 (Rel. 46, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Taste receptor type 1 member 3 precursor (Sweet taste receptor T1R3).
GN   Name=TAS1R3; Synonyms=T1R3;
OS   Gorilla gorilla gorilla (Lowland gorilla).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Gorilla.
OX   NCBI_TaxID=9595;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Li X., Reed D.R., Tordoff M.G., Bachmanov A.A., Mascioli K.J., Bak G.,
RA   Li W., Beauchamp G.K.;
RT   "Analysis of the T1R taste receptor genes in primates.";
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Putative taste receptor. TAS1R1/TAS1R3 responds to the
CC       umami taste stimulus (the taste of monosodium glutamate).
CC       TAS1R2/TAS1R3 recognizes diverse natural and synthetic sweeteners.
CC       TAS1R3 is essential for the recognition and response to the
CC       disaccharide trehalose (By similarity). Sequence differences
CC       within and between species can significantly influence the
CC       selectivity and specificity of taste responses (By similarity).
CC   -!- SUBUNIT: Forms homodimers or heterodimers with TAS1R1 and TAS1R2
CC       (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to the G-protein coupled receptor 3 family.
CC       TAS1R subfamily.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; AF545574; AAQ11897.1; -; Genomic_DNA.
DR   InterPro; IPR001828; ANF_rcpt_lig_bd.
DR   InterPro; IPR000068; Ca_sens_receptor.
DR   InterPro; IPR000337; GPCR_Mgr.
DR   InterPro; IPR011500; NCD3G_GPCR.
DR   Pfam; PF00003; 7tm_3; 1.
DR   Pfam; PF01094; ANF_receptor; 1.
DR   Pfam; PF07562; NCD3G; 1.
DR   PRINTS; PR00592; CASENSINGR.
DR   PRINTS; PR00248; GPCRMGR.
DR   PROSITE; PS00979; G_PROTEIN_RECEP_F3_1; FALSE_NEG.
DR   PROSITE; PS00980; G_PROTEIN_RECEP_F3_2; 1.
DR   PROSITE; PS00981; G_PROTEIN_RECEP_F3_3; FALSE_NEG.
DR   PROSITE; PS50259; G_PROTEIN_RECEP_F3_4; 1.
KW   G-protein coupled receptor; Glycoprotein; Receptor; Signal;
KW   Transducer; Transmembrane.
FT   SIGNAL        1     20       Potential.
FT   CHAIN        21    852       Taste receptor type 1 member 3.
FT   TOPO_DOM     21    570       Extracellular (Potential).
FT   TRANSMEM    571    591       1 (Potential).
FT   TOPO_DOM    592    603       Cytoplasmic (Potential).
FT   TRANSMEM    604    624       2 (Potential).
FT   TOPO_DOM    625    639       Extracellular (Potential).
FT   TRANSMEM    640    660       3 (Potential).
FT   TOPO_DOM    661    682       Cytoplasmic (Potential).
FT   TRANSMEM    683    703       4 (Potential).
FT   TOPO_DOM    704    729       Extracellular (Potential).
FT   TRANSMEM    730    750       5 (Potential).
FT   TOPO_DOM    751    762       Cytoplasmic (Potential).
FT   TRANSMEM    763    783       6 (Potential).
FT   TOPO_DOM    784    791       Extracellular (Potential).
FT   TRANSMEM    792    812       7 (Potential).
FT   TOPO_DOM    813    852       Cytoplasmic (Potential).
FT   CARBOHYD     85     85       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    130    130       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    264    264       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    285    285       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    380    380       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    411    411       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    432    432       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    475    475       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   852 AA;  93362 MW;  13F271103EB39F45 CRC64;

  Query Match             98.1%;  Score 4439;  DB 1;  Length 852;
  Best Local Similarity   98.1%;  Pred. No. 6.9e-314;
  Matches  836;  Conservative    6;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              |||||||||||||| || ||||||||||||||:|||||||||||||| |||||||||
Db          1 MLGPAVLGLSLWALLQPGAGAPLCLSQQLRMKGDYMLGGLFPLGEAEEAGFRSRTRPSSP 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALAADPAFCSALGEREQG 360

Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDP 420

Qy        421 VKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480
              |||||||||||||| ||| ||||||||||||||||||||||||||||||||||||||||
Db        421 VKPWQLLENMYNLTFHAGGLMLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRT 480

Qy        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTF 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||: ||
Db        481 ERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDVTCTS 540

Qy        541 CGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGQDEWSPERSTRCFHRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQ 600

Qy        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660
              |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        601 ASGGPLACFGLVCLGLVCLSVLLFPGQPSPAQCLAQQPLSHLPLTGCLSTLFLQAAEIFV 660

Qy        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALV 720

Qy        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVP 780
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        721 HCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGRYNRARGLTFAMLAYFITWVSFVP 780

Qy        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGGGPGDAQGQ 840
              |||||||||||||||||||||||||||||||||||:|||||||||||||||||||||:
Db        781 LLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLIRQPGLNTPEFFLGGGPGDAQGR 840
```

```
Qy        841 NDGNTGNQGKHE 852
              |||:||||||||
Db        841 NDGDTGNQGKHE 852


RESULT 4
TS1R3_RAT
ID   TS1R3_RAT      STANDARD;      PRT;   858 AA.
AC   Q923K1;
DT   01-FEB-2005 (Rel. 46, Created)
DT   01-FEB-2005 (Rel. 46, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Taste receptor type 1 member 3 precursor (Sweet taste receptor T1R3).
GN   Name=Tas1r3; Synonyms=T1r3, Tr3;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muroidea; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Wistar; TISSUE=Circumvallate papillae;
RX   MEDLINE=21400448; PubMed=11509186; DOI=10.1016/S0092-8674(01)00451-2;
RA   Nelson G., Hoon M.A., Chandrashekar J., Zhang Y., Ryba N.J.P.,
RA   Zuker C.S.;
RT   "Mammalian sweet taste receptors.";
RL   Cell 106:381-390(2001).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Wistar; TISSUE=Circumvallate papillae;
RX   MEDLINE=21927605; PubMed=11917125; DOI=10.1073/pnas.072090199;
RA   Li X., Staszewski L., Xu H., Durick K., Zoller M., Adler E.;
RT   "Human receptors for sweet and umami taste.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:4692-4696(2002).
CC   -!- FUNCTION: Putative taste receptor. TAS1R1/TAS1R3 responds to the
CC       umami taste stimulus (the taste of monosodium glutamate) and also
CC       to most of the 20 standard L-amino acids, but not to their D-
CC       enantiomers or other compounds. TAS1R2/TAS1R3 recognizes diverse
CC       natural and synthetic sweeteners. TAS1R3 is essential for the
CC       recognition and response to the disaccharide trehalose. Sequence
CC       differences within and between species can significantly influence
CC       the selectivity and specificity of taste responses (By
CC       similarity).
CC   -!- SUBUNIT: Forms homodimers or heterodimers with TAS1R1 and TAS1R2
CC       (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to the G-protein coupled receptor 3 family.
CC       TAS1R subfamily.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; AY032620; AAK51601.1; -; mRNA.
DR   EMBL; AF456324; AAM10636.1; -; mRNA.
DR   Ensembl; ENSRNOG00000019589; Rattus norvegicus.
DR   RGD; 620539; Tas1r3.
DR   GO; GO:0016021; C:integral to membrane; IC.
DR   GO; GO:0008527; F:taste receptor activity; IDA.
DR   GO; GO:0050916; P:perception of sweet taste; IDA.
DR   GO; GO:0050917; P:perception of umami taste; IDA.
DR   InterPro; IPR001828; ANF_rcpt_lig_bd.
DR   InterPro; IPR000068; Ca_sens_receptor.
DR   InterPro; IPR000337; GPCR_Mgr.
DR   InterPro; IPR011500; NCD3G_GPCR.
DR   Pfam; PF00003; 7tm_3; 1.
DR   Pfam; PF01094; ANF_receptor; 1.
DR   Pfam; PF07562; NCD3G; 1.
DR   PRINTS; PR00592; CASENSINGR.
DR   PRINTS; PR00248; GPCRMGR.
DR   PROSITE; PS00979; G_PROTEIN_RECEP_F3_1; FALSE_NEG.
DR   PROSITE; PS00980; G_PROTEIN_RECEP_F3_2; 1.
DR   PROSITE; PS00981; G_PROTEIN_RECEP_F3_3; FALSE_NEG.
DR   PROSITE; PS50259; G_PROTEIN_RECEP_F3_4; 1.
KW   G-protein coupled receptor; Glycoprotein; Receptor;
KW   Sensory transduction; Signal; Taste; Transducer; Transmembrane.
FT   SIGNAL        1     20       Potential.
FT   CHAIN        21    858       Taste receptor type 1 member 3.
FT   TOPO_DOM     21    575       Extracellular (Potential).
FT   TRANSMEM    576    596       1 (Potential).
FT   TOPO_DOM    597    610       Cytoplasmic (Potential).
FT   TRANSMEM    611    631       2 (Potential).
FT   TOPO_DOM    632    644       Extracellular (Potential).
FT   TRANSMEM    645    665       3 (Potential).
FT   TOPO_DOM    666    687       Cytoplasmic (Potential).
FT   TRANSMEM    688    708       4 (Potential).
FT   TOPO_DOM    709    735       Extracellular (Potential).
FT   TRANSMEM    736    756       5 (Potential).
FT   TOPO_DOM    757    767       Cytoplasmic (Potential).
FT   TRANSMEM    768    788       6 (Potential).
FT   TOPO_DOM    789    796       Extracellular (Potential).
FT   TRANSMEM    797    817       7 (Potential).
FT   TOPO_DOM    818    858       Cytoplasmic (Potential).
FT   CARBOHYD     85     85       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    130    130       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    203    203       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    264    264       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    379    379       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    387    387       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    418    418       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    439    439       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    482    482       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   858 AA;  94846 MW;  98890DAF75973B80 CRC64;

  Query Match             71.9%;  Score 3255;  DB 1;  Length 858;
  Best Local Similarity   72.9%;  Pred. No. 9.1e-228;
  Matches  623;  Conservative   74;  Mismatches  144;  Indels   14;  Gaps    4;

Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              | | |:||||| | |   | |: |||||| : :|||:||||||||  ||| |  ||:|:
Db          1 MPGLAILGLSLAAFLELGMGSSLCLSQQFKAQGDYILGGLFPLGTTEEATLNQRTQPNGI 60

Qy         61 VCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKA 120
              :|||||  ||  |:|||||||||||| | |||||||||||||||||||||| |||||||:||
Db         61 LCTRFSPLGLFLAMAMKMAVEEINNGSALLPGLRLGYDLFDTCSEPVVTMKPSLMFMAKV 120

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              ||: |||||||||||||||||||||||||::|||||||||||||||| |||: || ||||
Db        121 GSQSIAAYCNYTQYQPRVLAVIGPHSSELALITGKFFSFFLMPQVSYSASMDRLSDRETF 180

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              |||||||||||||  |   ||| | ||||||||||:|||:|||||| || :||||||||
Db        181 PSFFRTVPSDRVQLQAVVTLLQNFSWNWVAALGSDDDYGREGLSIFSGLANSRGICIAHE 240

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              ||||       :|||| ||| ||||| ||||:||||  | ::||:|||   |||||||||
Db        241 GLVPQHDTSGQQLGKVVDVLRQVNQSKVQVVVLFASARAVYSLFSYSILHDLSPKVWVAS 300

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQG 360
              |:|||||||| || :|::|||||||||||| | ||  ||:| |||| || ||::| : |
Db        301 ESWLTSDLVMTLPNIARVGTVLGFLQRGALLPEFSHYVETRLALAADPTFCASL-KAELD 359

Qy        361 LEEDVVGQRCPQCDCITLQNVSAGLN--------HHQTFSVYAAVYSVAQALHNTLQCNA 412
              ||| |:| || ||| | |||:|:||         ||| |: ||||||||||||||||||
Db        360 LEERVMGPRCSQCDYIMLQNLSSGLMQNLSAGQLHHQIFATYAAVYSVAQALHNTLQCNV 419

Qy        413 SGCPAQDPVKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVG 472
              | |   :||:|||||||||::|    | |:||: |:|||||||:|||   | || ||
Db        420 SHCHTSEPVQPWQLLENMYNMSFRARDLTLQFDAKGSVDMEYDLKMWVWQSPTPVLHTVG 479

Qy        473 RFNGSLRTERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQN 532
              |||:|: :  |: |    || |||:|||||::|||||||||||||||||:|||||::
Db        480 TFNGTLQLQHSKMYW--PGNQVPVSQCSRQCKDGQVRRVKGFHSCCYDCVDCKAGSYRKH 537
```

```
Qy        533 PDDIACTFCGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVH 592
              |||  || ||:|:||||:|| |  || :|||||||||| ||||| | ||| |||||||||
Db        538 PDDFTCTPCGKDQWSPEKSTTCLPRRPKFLAWGEPAVLSLLLLLCLVLGLTLAALGLFVH 597

Qy        593 HRDSPLVQASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLF 652
              : |||||||||| | ||||:|||| |||||||||:|  | ||||||::||||||||||||
Db        598 YWDSPLVQASGGSLFCFGLICLGLFCLSVLLFPGRPRSASCLAQQPMAHLPLTGCLSTLF 657

Qy        653 LQAAEIFVESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWH 712
              |||||||||||||||||: |   ||||||||||||| ||| ||| |||:|||||||||||
Db        658 LQAAEIFVESELPLSWANWLCSYLRGPWAWLVVLLATLVEAALCAWYLMAFPPEVVTDWQ 717

Qy        713 MLPTEALVHCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYF 772
              :|||| | ||| ||||| || | ||| |||||||||||:|||| ||||||||||||||||
Db        718 VLPTEVLEHCRMRSWVSLGLVHITNAVLAFLCFLGTFLVQSQPGRYNRARGLTFAMLAYF 777

Qy        773 ITWVSFVPLLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLGG 832
              | |||||||||||||  :|||||||:| | ||||| ||||:||:|:  | ||| |||||
Db        778 IIWVSFVPLLANVQVAYQPAVQMGAILFCALGILATFHLPKCYVLLWLPELNTQEFFLGR 837

Qy        833 GPGDAQGQNDGNTGN 847
              | :|   :|||:|:
Db        838 SPKEA---SDGNSGS 849
```

```
RESULT 5
TS1R3_MOUSE
ID   TS1R3_MOUSE    STANDARD;      PRT;   858 AA.
AC   Q925D8; Q91VA4; Q923K0; Q925A4; Q925D9;
DT   01-FEB-2005 (Rel. 46, Created)
DT   01-FEB-2005 (Rel. 46, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Taste receptor type 1 member 3 precursor (Sweet taste receptor T1R3)
DE   (Saccharin preference protein).
GN   Name=Tas1r3; Synonyms=Sac, T1r3, Tr3;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muroidea; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE, TISSUE SPECIFICITY, AND VARIANTS ALA-55; THR-60;
RP   LEU-61; GLN-371 AND THR-706.
RC   STRAIN=C57BL/6N, and C57BL/6NCr;
RC   TISSUE=Brain, and Circumvallate papillae;
RX   MEDLINE=21222875; PubMed=11322794; DOI=10.1006/bbrc.2001.4760;
RA   Kitagawa M., Kusakabe Y., Miura H., Ninomiya Y., Hino A.;
RT   "Molecular genetic identification of a candidate receptor gene for
RT   sweet taste.";
RL   Biochem. Biophys. Res. Commun. 283:236-242(2001).
RN   [2]
RP   NUCLEOTIDE SEQUENCE, TISSUE SPECIFICITY, AND VARIANTS ALA-55; THR-60;
RP   LEU-61; CYS-261; GLN-371; SER-692; THR-706 AND GLU-855.
RC   STRAIN=C57BL/6J, DBA/2J, and SWR/J;
RC   TISSUE=Circumvallate papillae, and Foliate papillae;
RX   MEDLINE=21219400; PubMed=11319557; DOI=10.1038/87440;
RA   Montmayeur J.-P., Liberles S.D., Matsunami H., Buck L.B.;
RT   "A candidate taste receptor gene near a sweet taste locus.";
RL   Nat. Neurosci. 4:492-498(2001).
RN   [3]
RP   NUCLEOTIDE SEQUENCE, AND VARIANTS ALA-55 AND THR-60.
RC   STRAIN=129/J, AKR/J, BALB/cByJ, C3H/HeJ, C57BL/6ByJ, C57L/J, CAST/Ei,
RC   CBA/J, DBA/2J, IS/CamEi, SEA/GnJ, ST/bJ, and SWR/J;
RX   MEDLINE=21439023; PubMed=11555487; DOI=10.1093/chemse/26.7.925;
RA   Bachmanov A.A., Li X., Reed D.R., Ohmen J.D., Li S., Chen Z.,
RA   Tordoff M.G., de Jong P.J., Wu C., West D.B., Chatterjee A.,
RA   Ross D.A., Beauchamp G.K.;
RT   "Positional cloning of the mouse saccharin preference (Sac) locus.";
RL   Chem. Senses 26:925-933(2001).
RN   [4]
RP   NUCLEOTIDE SEQUENCE, AND VARIANTS ALA-55; THR-60; LEU-61; CYS-261;
RP   GLN-371; SER-692; THR-706 AND GLU-855.
RC   STRAIN=129/Sv, AKR/J, BALB/c, C3H/HeJ, C57BL/6, C57L/J, DBA/1LacJ,
RC   DBA/2J, FVB/N, ST/bJ, and SWR/J;
RX   MEDLINE=21400448; PubMed=11509186; DOI=10.1016/S0092-8674(01)00451-2;
RA   Nelson G., Hoon M.A., Chandrashekar J., Zhang Y., Ryba N.J.P.,
RA   Zuker C.S.;
RT   "Mammalian sweet taste receptors.";
RL   Cell 106:381-390(2001).
RN   [5]
RP   NUCLEOTIDE SEQUENCE, TISSUE SPECIFICITY, AND VARIANTS ALA-55; THR-60;
RP   LEU-61; GLN-371; THR-706 AND GLU-855.
RC   STRAIN=129/Sv, and C57BL/6;
RX   PubMed=11331418;
RA   Sainz E., Korley J.N., Battey J.F., Sullivan S.L.;
RT   "Identification of a novel member of the T1R family of putative taste
RT   receptors.";
RL   J. Neurochem. 77:896-903(2001).
RN   [6]
RP   NUCLEOTIDE SEQUENCE, AND VARIANTS ALA-55; THR-60 AND LEU-61.
RC   STRAIN=129/SvEv, and C57BL/6J;
RX   MEDLINE=21635547; PubMed=11773963;
RA   Li X., Bachmanov A.A., Li S., Chen Z., Tordoff M.G., Beauchamp G.K.,
RA   de Jong P.J., Wu C., Chen L., West D.B., Ross D.A., Ohmen J.D.,
RA   Reed D.R.;
RT   "Genetic, physical, and comparative map of the subtelomeric region of
RT   mouse Chromosome 4.";
RL   Mamm. Genome 13:5-19(2002).
RN   [7]
RP   NUCLEOTIDE SEQUENCE, FUNCTION, AND VARIANT ASN-706.
RC   STRAIN=C57BL/6;
RX   PubMed=12892531; DOI=10.1290/1543-706X(2003)039<0080:TRTIAE>2.0.CO;2;
RA   Ariyasu T., Matsumoto S., Kyono F., Hanaya T., Arai S., Ikeda M.,
RA   Kurimoto M.;
RT   "Taste receptor T1R3 is an essential molecule for the cellular
RT   recognition of the disaccharide trehalose.";
RL   In Vitro Cell. Dev. Biol. Anim. 39:80-88(2003).
RN   [8]
RP   NUCLEOTIDE SEQUENCE, VARIANTS ALA-55; THR-60; LEU-61; CYS-261;
RP   GLN-371; SER-692 AND THR-706, AND TISSUE SPECIFICITY.
RC   STRAIN=129/SvEv, BALB/c, C3H/HeJ, C57BL/6, DBA/2J, FVB/N, ST/bJ, and
RC   SWR/J;
RX   MEDLINE=21225287; PubMed=11326277; DOI=10.1038/88270;
RA   Max M., Shanker Y.G., Huang L., Rong M., Liu Z., Campagne F.,
RA   Weinstein H., Damak S., Margolskee R.F.;
RT   "Tas1r3, encoding a new candidate taste receptor, is allelic to the
RT   sweet responsiveness locus Sac.";
RL   Nat. Genet. 28:58-63(2001).
RN   [9]
RP   FUNCTION.
RX   PubMed=11894099; DOI=10.1038/nature726;
RA   Nelson G., Chandrashekar J., Hoon M.A., Feng L., Zhao G., Ryba N.J.P.,
RA   Zuker C.S.;
RT   "An amino-acid taste receptor.";
RL   Nature 416:199-202(2002).
RN   [10]
RP   SWEET AND UMAMI TASTES EXCLUSIVELY MEDIATED BY TAS1R RECEPTORS.
RX   PubMed=14636554; DOI=10.1016/S0092-8674(03)00844-4;
RA   Zhao G.Q., Zhang Y., Hoon M.A., Chandrashekar J., Erlenbach I.,
RA   Ryba N.J.P., Zuker C.S.;
RT   "The receptors for mammalian sweet and umami taste.";
RL   Cell 115:255-266(2003).
RN   [11]
RP   TAS1R3 INDEPENDENT SWEET- AND UMAMI-RESPONSIVE RECEPTORS.
RX   PubMed=12869700; DOI=10.1126/science.1087155;
RA   Damak S., Rong M., Yasumatsu K., Kokrashvili Z., Varadarajan V.,
RA   Zou S., Jiang P., Ninomiya Y., Margolskee R.F.;
RT   "Detection of sweet and umami taste in the absence of taste receptor
RT   T1r3.";
RL   Science 301:850-853(2003).
RN   [12]
RP   VARIANTS ALA-55; THR-60; LEU-61; CYS-261; GLN-371; SER-692; ASN-706;
RP   THR-706 AND GLU-855.
RC   STRAIN=129/J, AKR/J, C57BL/6, CAST/Ei, DBA/2J, and SWR/J;
RX   PubMed=14749438; DOI=10.1523/JNEUROSCI.1374-03.2004;
```

```
RA   Reed D.R., Li S., Li X., Huang L., Tordoff M.G., Starling-Roney R.,
RA   Taniguchi K., West D.B., Ohmen J.D., Beauchamp G.K., Bachmanov A.A.;
RT   "Polymorphisms in the taste receptor gene (Tas1r3) region are
RT   associated with saccharin preference in 30 mouse strains.";
RL   J. Neurosci. 24:938-946(2004).
RN   [13]
RP   ALLELIC VARIATION RESPONSE TO SWEETENERS.
RC   STRAIN=129/J, and C57BL/ByJ;
RX   PubMed=14999080; DOI=10.1523/JNEUROSCI.4439-03.2004;
RA   Inoue M., Reed D.R., Li X., Tordoff M.G., Beauchamp G.K.,
RA   Bachmanov A.A.;
RT   "Allelic variation of the Tas1r3 taste receptor gene selectively
RT   affects behavioral and neural taste responses to sweeteners in the F2
RT   hybrids between C57BL/6ByJ and 129P3/J mice.";
RL   J. Neurosci. 24:2296-2303(2004).
CC   -!- FUNCTION: Putative taste receptor. TAS1R1/TAS1R3 responds to the
CC       umami taste stimulus (the taste of monosodium glutamate) and also
CC       to most of the 20 standard L-amino acids, but not to their D-
CC       enantiomers or other compounds. TAS1R2/TAS1R3 recognizes diverse
CC       natural and synthetic sweeteners. TAS1R3 is essential for the
CC       recognition and response to the disaccharide trehalose. Sequence
CC       differences within and between species can significantly influence
CC       the selectivity and specificity of taste responses.
CC   -!- SUBUNIT: Forms homodimers or heterodimers with TAS1R1 and TAS1R2.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Expressed in circumvallate, foliate and
CC       fungiform taste papillae as well as in taste buds on the palate.
CC       Also expressed in testis. Not expressed in brain, heart, kidney,
CC       liver or spleen. The topographic distribution in various taste
CC       papillae is different from those of other T1R members.
CC   -!- PTM: The Thr-60 variant is predicted to introduce a novel N-linked
CC       glycosylation site at Asn-58. The addition of even a short
CC       carbohydrate group at Asn-58 is predicted to disrupt one of the
CC       contact surfaces required for stability of a dimer. Therefore the
CC       predicted N-linked glycosyl group at Asn-58 would be expected to
CC       preclude from forming homodimers or heterodimers.
CC   -!- SIMILARITY: Belongs to the G-protein coupled receptor 3 family.
CC       TAS1R subfamily.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; AB049994; BAB47181.1; -; mRNA.
DR   EMBL; AB055708; BAB62852.1; -; Genomic_DNA.
DR   EMBL; AF337039; AAK39436.1; -; mRNA.
DR   EMBL; AF311386; AAL08425.1; -; mRNA.
DR   EMBL; AY032621; AAK51602.1; -; mRNA.
DR   EMBL; AY026318; AAK01937.1; -; Genomic_DNA.
DR   EMBL; AF368024; AAK55536.1; -; Genomic_DNA.
DR   EMBL; AF368025; AAK55537.1; -; Genomic_DNA.
DR   MGI; MGI:1933547; Tas1r3.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0004930; F:G-protein coupled receptor activity; TAS.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; TAS.
DR   InterPro; IPR001828; ANF_rcpt_lig_bd.
DR   InterPro; IPR000068; Ca_sens_receptor.
DR   InterPro; IPR000337; GPCR_Mgr.
DR   InterPro; IPR011500; NCD3G_GPCR.
DR   Pfam; PF00003; 7tm_3; 1.
DR   Pfam; PF01094; ANF_receptor; 1.
DR   Pfam; PF07562; NCD3G; 1.
DR   PRINTS; PR00592; CASENSINGR.
DR   PRINTS; PR00248; GPCRMGR.
DR   PROSITE; PS00979; G_PROTEIN_RECEP_F3_1; FALSE_NEG.
DR   PROSITE; PS00980; G_PROTEIN_RECEP_F3_2; 1.
DR   PROSITE; PS00981; G_PROTEIN_RECEP_F3_3; FALSE_NEG.
DR   PROSITE; PS50259; G_PROTEIN_RECEP_F3_4; 1.
KW   G-protein coupled receptor; Glycoprotein; Polymorphism; Receptor;
KW   Sensory transduction; Signal; Taste; Transducer; Transmembrane.
```

```
FT   SIGNAL       1    20       Potential.
FT   CHAIN       21   858       Taste receptor type 1 member 3.
FT   TOPO_DOM    21   572       Extracellular (Potential).
FT   TRANSMEM   573   593       1 (Potential).
FT   TOPO_DOM   594   610       Cytoplasmic (Potential).
FT   TRANSMEM   611   631       2 (Potential).
FT   TOPO_DOM   632   644       Extracellular (Potential).
FT   TRANSMEM   645   665       3 (Potential).
FT   TOPO_DOM   666   687       Cytoplasmic (Potential).

  Query Match             71.4%;  Score 3229;  DB 1;  Length 858;
  Best Local Similarity   72.5%;  Pred. No. 7.1e-226;
  Matches  621;  Conservative   67;  Mismatches  147;  Indels   22;  Gaps    4;

Qy          5 AVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSPVCTR 64
              |::|||| | |  | || |||||| : :|||:||||||||  ||| |  ||:|:|  | |
Db          5 AIMGLSLAAFLELGMGASLCLSQQFKAQGDYILGGLFPLGSTEEATLNQRTQPNSIPCNR 64

Qy         65 FSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKAGSRD 124
              ||  ||  |:|||||||||||| | ||||||||||||||||||||| || ||||||| ||:
Db         65 FSPLGLFLAMAMKMAVEEINNGSALLPGLRLGYDLFDTCSEPVVTMKSSLMFLAKVGSQS 124

Qy        125 IAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETFPSFF 184
              ||||||||||||||||||||||||::|||||||||||||| |||: || ||||||||
Db        125 IAAYCNYTQYQPRVLAVIGPHSSELALITGKFFSFFLMPQVSYSASMDRLSDRETFPSFF 184

Qy        185 RTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHEGLVP 244
              |||||||||| |   ||| | |||||||||||:|||:|||||:|| ||||||||||||
Db        185 RTVPSDRVQLQAVVTLLQNFSWNWVAALGSDDDYGREGLSIFSSLANARGICIAHEGLVP 244

Qy        245 LPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVASEAWL 304
                  :|||| ||| ||||| ||||:|||| | ::||:|||   ||||||||||:||
Db        245 QHDTSGQQLGKVLDVLRQVNQSKVQVVVLFASARAVYSLFSYSIHHGLSPKVWVASESWL 304

Qy        305 TSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQGLEED 364
              |||||| || :|::|||||||||||| | ||  ||:|||||| ||||||::|   |  |||
Db        305 TSDLVMTLPNIARVGTVLGFLQRGALLPEFSHYVETHLALAADPAFCASL-NAELDLEEH 363

Qy        365 VVGQRCPQCDCITLQNVSAGL--------NHHQTFSVYAAVYSVAQALHNTLQCNASGCP 416
              |:|||||:|| | |||:|:||         ||| |: |||||||||||||||||| | |
Db        364 VMGQRCPRCDDIMLQNLSSGLLQNLSAGQLHHQIFATYAAVYSVAQALHNTLQCNVSHCH 423

Qy        417 AQDPVKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNG 476
              : | |||||||||::||   | |:||: ||||||||||:||||   | || || |||
Db        424 VSEHVLPWQLLENMYNMSFHARDLTLQFDAEGNVDMEYDLKMWVWQSPTPVLHTVGTFNG 483

Qy        477 SLRTERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDI 536
              :|: :: |: |    || |||:|||||::|||||||||||||||||:|||||::|||
Db        484 TLQLQQSKMYW--PGNQVPVSQCSRQCKDGQVRRVKGFHSCCYDCVDCKAGSYRKHPDDF 541

Qy        537 ACTFCGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDS 596
              || | ||:||||:|| |  || :|||||| || ||||| | ||| |||||| ||| ||
Db        542 TCTPCNQDQWSPEKSTACLPRRPKFLAWGEPVVLSLLLLLCLVLGLALAALGLSVHHWDS 601

Qy        597 PLVQASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAA 656
              ||||||||   ||||:|||| |||||||||:|| | |||||||::||||||||||||||
Db        602 PLVQASGGSQFCFGLICLGLFCLSVLLFPGRPSSASCLAQQPMAHLPLTGCLSTLFLQAA 661

Qy        657 EIFVESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPT 716
              | |||||||||||: |   ||| |||||||||   || ||| |||:|||||||||| :|||
Db        662 ETFVESELPLSWANWLCSYLRGLWAWLVVLLATFVEAALCAWYLIAFPPEVVTDWSVLPT 721

Qy        717 EALVHCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWV 776
              | | ||  ||||| || | ||| ||||||||||||:|||| ||||||||||||||||||
Db        722 EVLEHCHVRSWVSLGLVHITNAMLAFLCFLGTFLVQSQPGRYNRARGLTFAMLAYFITWV 781

Qy        777 SFVPLLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLG----- 831
              ||||||||||  :|||||||:|:| |||| |||||:||:|:  | ||| |||||
Db        782 SFVPLLANVQVAYQPAVQMGAILVCALGILVTFHLPKCYVLLWLPKLNTQEFFLGRNAKK 841

Qy        832 ------GGPGDAQGQND 842
                     ||  ||| |:
Db        842 AADENSGGGEAAQGHNE 858
```

```
RESULT 6
Q4RX45_TETNG
ID   Q4RX45_TETNG PRELIMINARY;      PRT;   829 AA.
AC   Q4RX45;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Chromosome 11 SCAF14979, whole genome shotgun sequence.
DE   (Fragment).
GN   ORFNames=GSTENG00027570001;
OS   Tetraodon nigroviridis (Green puffer).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Tetraodon.
OX   NCBI_TaxID=99883;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA   Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA   Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA   Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA   Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA   Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA   Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA   Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA   Kellis M., Volff JN., Guigo R., Zody M.C., Mesirov J.,
RA   Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA   Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA   Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT   "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT   the early vertebrate proto-karyotype.";
RL   Nature 431:946-957(2004).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RG   Genoscope; Whitehead Institute Centre for Genome Research;
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAE01014979; CAG07037.1; -; Genomic_DNA.
FT   NON_TER     829    829
SQ   SEQUENCE   829 AA;  92350 MW;  D5C8FA48E72B86BF CRC64;

  Query Match             32.8%;  Score 1483.5;  DB 2;  Length 829;
  Best Local Similarity   39.2%;  Pred. No. 6.1e-99;
  Matches  320;  Conservative  139;  Mismatches  320;  Indels   37;  Gaps   13;

Qy         31 MKGDYVLGGLFPLGEAEEAGLRSRTRPSSPVCTRFSSNGLLWALAMKMAVEEINNKSDLL 90
              : || :||||||: :   : |  || |    | |  | ||  |: || ||:|||    ||
Db         34 LPGDIMLGGLFPINQL-TSNLSQRTEPDQITCDRIDSFGLGMAIVMKYAVDEINANQFLL 92

Qy         91 PGLRLGYDLFDTCSEPVVAMKPSLMFLAKAGSRDIAAYCNYTQYQPRVLAVIGPHSSELA 150
              ||:||||:: ||| :  | ::|:: :|      :: | |||| |: |: |||||::|||
Db         93 PGIRLGYEIHDTCRQSAVIIRPTIYYLRAKHDNNLIAQCNYTNYETRISAVIGPNNSELV 152

Qy        151 MVTGKFFSFFLMPQVSYGASMELLSARETFPSFFRTVPSDRVQLTAAAELLQEFGWNWVA 210
              | ||   ||||||:||||: :  |    :||||||||||: |: |   || || |||||
Db        153 SVIGKLLGFFLMPQISYGATSDKFSDNILYPSFFRTVPSDKGQVEAMVLLLLEFNWNWVA 212

Qy        211 ALGSDDEYGRQGLSIFSALAAARGICIAHEGLVPLPRADDSRLGKVQDVLHQVNQSSVQV 270
              :|||:|||::|:  || :|| | ||:|::||:|:    :     |: ::  :| :   |
Db        213 VVGSDEEYGQRGVQDFSKVAANRSICVAYQGLIPVYTDPEP---VVKTIISNINSTKANV 269

Qy        271 VLLFASVHAAHALFNYSISSRLSPKVWVASEAWLTSDLVMGLPGMAQMGTVLGF------ 324
              |::|:  : |   |   | : |   ||: | ||  |: |  || : :||:|||
Db        270 VIVFSLPNQAEIFFQEVIRTELK-GVWIGSTAWTISERVFLLPNLHTVGTILGFAFTTQS 328

Qy        325 --LQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQGLEEDVVGQRCPQCDCITLQNVS 382
                | | |  ||    |    |    |              : | :      || |  ::  |:|
Db        329 LDLLR-AYTHELLTKVSEEHAHTPPPV---------ENLND--YSNSCPHCWNLSPANIS 376

Qy        383 AGLN---HHQTFSVYAAVYSVAQALHNTLQCNASGCPAQDPVK--PWQLLENMYNLTFHV 437
                       | ||||:|||||||| ||||:: |      |:  ||:||: : | |  :
Db        377 LDAEVAIQRSAFRVYAAIYSVAQALHNLLQCNSTACMKTSEVQIYPWKLLKALRNTTVDI 436

Qy        438 GGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRTERLKIRWHTSDNQKPVS 497
              |  |:|| :|| :: | :   : : |      :|| ||  |  ::   :||| ::: | |
Db        437 NGTMLKFDRNGNPNIGYSVVELILKNSSLEFLEVGSFNKILDIKKSLFKWHTENSEVPES 496

Qy        498 RCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTFCGQDEWSPERSTRCFRR 557
              ||  | |||| |||||||||:||:||  |:|::  ||| || |   :|| :|||||
Db        497 TCSAACGEGQVHRVKGFHSCCFDCIDCLPGTYQREKDDIQCTPCPPRQWSLKRSTRCTDP 556

Qy        558 RSRFLAWGEPAVLLLLLLSLALGLVLAALGLFVHHRDSPLVQASGGPLACFGLVCLGLV 617
                :|:|  |  ||| | : | : |  | : ||: || : || |||| |:   :| |||:
Db        557 TYDYLSWDTPEALLLTLAIVLVVLLKGAVVVLFLKHRGTVLVAASGGTLSF--VVLLGLM 614

Qy        618 --CLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFVESELPLSWADRLSGC 675
                |||:||| |||    |  | ||  :  |  |  :   : :||  || |   |  |
Db        615 GACLSLLLFLGQPGDTVCRLQLPLVSIFQTVPLCIIMSISLQIFFVSEFPTLAASYLH-V 673

Qy        676 LRGPWAWLVVLLAMLVEVALCTWYLVAFP--PEVVTDWHMLPTEALVHCRTRSWVSFGLA 733
              ||||  ||::|:   |:  :| |::    |   | : :  :    : : |   |   |||
Db        674 LRGPGTWLLLLICCAVQAGICGWFVQEGPSLSEYLANRRIDFVRSFLACPVSSLSGFGLM 733

Qy        734 HATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVPLLANVQVVLRPAV 793
                  :| : |: ||:        || || :||: | | | ||:|:|:   ::   :  |
Db        734 QGLITVMALISFMCTFMATKPLHQYNLARDITFSSLIYCIIWVTFIPIYVGLEEKFKSIV 793

Qy        794 QMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFF 829
                :  :|   ||::| :::|:||||:|:| ||| : |
Db        794 YVSFILTSNLGMVAMYYIPKCYLLLRKPELNTADHF 829


RESULT 7
Q4RX46_TETNG
ID   Q4RX46_TETNG PRELIMINARY;      PRT;   857 AA.
AC   Q4RX46;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Chromosome 11 SCAF14979, whole genome shotgun sequence.
DE   (Fragment).
GN   ORFNames=GSTENG00027569001;
OS   Tetraodon nigroviridis (Green puffer).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Tetraodon.
OX   NCBI_TaxID=99883;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA   Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA   Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA   Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA   Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA   Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA   Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA   Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA   Kellis M., Volff JN., Guigo R., Zody M.C., Mesirov J.,
RA   Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA   Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA   Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT   "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT   the early vertebrate proto-karyotype.";
RL   Nature 431:946-957(2004).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RG   Genoscope; Whitehead Institute Centre for Genome Research;
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
```

```
CC       preliminary data.
DR   EMBL; CAAE01014979; CAG07036.1; -; Genomic_DNA.
FT   NON_TER       1      1
SQ   SEQUENCE   857 AA;  96165 MW;  DC30B540EE7A104E CRC64;

  Query Match             27.2%;  Score 1232;  DB 2;  Length 857;
  Best Local Similarity   40.1%;  Pred. No. 1.3e-80;
  Matches  260;  Conservative  113;  Mismatches  245;  Indels   30;  Gaps    9;

Qy         64 RFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKAGSR 123
              |  : ||  |: ||  |:|||    ||||::|||:: ||| :  | ::|:: :||   :
Db          1 RIDTYGLGMAIVMKYTVDEINANQFLLPGIQLGYEIHDTCRQSAVIVRPTMYYLAAKSNN 60

Qy        124 DIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETFPSF 183
               ::  |||| |: || |||||: ||:  | ||   ||||||:||||: :  |    :|||
Db         61 MLSVECNYTNYETRVSAVIGPYGSEMVSVIGKLLGFFLMPQISYGATSDKFSDNILYPSF 120

Qy        184 FRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHEGLV 243
              ||||||: |: |   || || ||||| :|||:|||::|:  || :|| | ||:|::||:
Db        121 FRTVPSDKGQVEAMILLLLEFNWNWVAVVGSDEEYGQRGVQDFSKVAANRSICVAYQGLI 180

Qy        244 PLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVASEAW 303
              |:    :     |: ::  :| :   ||::|:  : |   |   | ::|   ||: | :|
Db        181 PVYTDPEP---VVKTIISNINSTKANVVIVFSLPNQAEIFFQEVIRTKLK-GVWIGSTSW 236

Qy        304 LTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQGLEE 363
               : |  || :  :||:| |       |      |    ::          |
Db        237 GIHNQVTSLPDIQTVGTILAFTDITQSLDHLSAYTHELLTKVSE----------EHAHTP 286

Qy        364 DVV------GQRCPQCDCITLQNVS---AGLNHHQTFSVYAAVYSVAQALHNTLQCNASG 414
               :|        ||:|  ::  |:|     :       ||||||:||||||| ||||::
Db        287 PMVQISAHPPNPCPKCWNLSPANISLVTQPVVQRTAFSVYAAIYSVAQALHNLLQCNSTA 346

Qy        415 CPAQDPVK--PWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVG 472
               |      ||  ||:||: : | : :: |  |:||::|| :: |:|  |||: |     ::|
Db        347 CMKTSEVKIYPWKLLKVLRNTSVNISGTQLKFDTNGNPNVGYNLVKWVWEESGVEFTEIG 406

Qy        473 RFNGSLRTERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQN 532
               |:  |         |||:::: | | || *| |||| |||||||||:||:||  |:|:::
Db        407 SFDEKLDINTSLFTWHTANSEVPESTCSAACGEGQVHRVKGFHSCCFDCIDCLPGTYQRD 466

Qy        533 PDDIACTFCGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLSLALGLVLAALGLFVH 592
               ||| || |   :|| :|||||      :|:|  |  ||| | : | : |  | :.||:
Db        467 ADDIQCTPCPPRQWSLKRSTRCTDPTYDYLSWDTPEALLLTLAIVLVVLLKGAVVVLFLK 526

Qy        593 HRDSPLVQASGGPLACFGLVCLGLV--CLSVLLFPGQPSPARCLAQQPLSHLPLTGCLST 650
              || : || |||| |:   :| |||:  |||:||| |||     |  | ||  :  | |
Db        527 HRGTVLVAASGGTLSF--VVLLGLMGACLSLLLFLGQPGDTVCRLQLPLVSIFQTVPLCI 584

Qy        651 LFLQAAEIFVESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTW 698
               :   : :||  || |   |  |   ||||  ||::|:   |:  :| |
Db        585 IMSISLQIFFVSEFPTLAASYLH-VLRGPGTWLLLLICCAVQAGICGW 631


RESULT 8
TS1R1_RAT
ID   TS1R1_RAT      STANDARD;      PRT;   840 AA.
AC   Q9Z0R8;
DT   01-FEB-2005 (Rel. 46, Created)
DT   01-FEB-2005 (Rel. 46, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Taste receptor type 1 member 1 precursor (G-protein coupled receptor
DE   70).
GN   Name=Tas1r1; Synonyms=Gpr70, T1r1, Tr1;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muroidea; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   NUCLEOTIDE SEQUENCE, AND TISSUE SPECIFICITY.
RC   STRAIN=Wistar; TISSUE=Circumvallate papillae;
RX   MEDLINE=99159821; PubMed=10052456; DOI=10.1016/S0092-8674(00)80658-3;
RA   Hoon M.A., Adler E., Lindemeier J., Battey J.F., Ryba N.J.P.,
RA   Zuker C.S.;
RT   "Putative mammalian taste receptors: a class of taste-specific GPCRs
RT   with distinct topographic selectivity.";
RL   Cell 96:541-551(1999).
CC   -!- FUNCTION: Putative taste receptor. TAS1R1/TAS1R3 responds to the
CC       umami taste stimulus (the taste of monosodium glutamate). Sequence
CC       differences within and between species can significantly influence
CC       the selectivity and specificity of taste responses.
CC   -!- SUBUNIT: Forms heterodimers with TAS1R3.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Expressed in many cells in the fungiform and
CC       geschmackstreifen taste buds.
CC   -!- SIMILARITY: Belongs to the G-protein coupled receptor 3 family.
CC       TAS1R subfamily.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; AF127389; AAD18069.1; -; mRNA.
DR   Ensembl; ENSRNOG00000009708; Rattus norvegicus.
DR   RGD; 61891; Tas1r1.
DR   GO; GO:0016021; C:integral to membrane; IC.
DR   GO; GO:0008527; F:taste receptor activity; IDA.
DR   GO; GO:0050917; P:perception of umami taste; IDA.
DR   InterPro; IPR001828; ANF_rcpt_lig_bd.
DR   InterPro; IPR000337; GPCR_Mgr.
DR   InterPro; IPR011500; NCD3G_GPCR.
DR   Pfam; PF00003; 7tm_3; 1.
DR   Pfam; PF01094; ANF_receptor; 1.
DR   Pfam; PF07562; NCD3G; 1.
DR   PRINTS; PR00248; GPCRMGR.
DR   PROSITE; PS00979; G_PROTEIN_RECEP_F3_1; FALSE_NEG.
DR   PROSITE; PS00980; G_PROTEIN_RECEP_F3_2; FALSE_NEG.
DR   PROSITE; PS00981; G_PROTEIN_RECEP_F3_3; FALSE_NEG.
DR   PROSITE; PS50259; G_PROTEIN_RECEP_F3_4; 1.
KW   G-protein coupled receptor; Glycoprotein; Receptor;
KW   Sensory transduction; Signal; Taste; Transducer; Transmembrane.
FT   SIGNAL        1     19       Potential.
FT   CHAIN        20    840       Taste receptor type 1 member 1.
FT   TOPO_DOM     20    566       Extracellular (Potential).
FT   TRANSMEM    567    587       Potential.
FT   TOPO_DOM    588    602       Cytoplasmic (Potential).
FT   TRANSMEM    603    623       Potential.
FT   TOPO_DOM    624    638       Extracellular (Potential).
FT   TRANSMEM    639    659       Potential.
FT   TOPO_DOM    660    679       Cytoplasmic (Potential).
FT   TRANSMEM    680    700       Potential.
FT   TOPO_DOM    701    724       Extracellular (Potential).
FT   TRANSMEM    725    745       Potential.
FT   TOPO_DOM    746    760       Cytoplasmic (Potential).
FT   TRANSMEM    761    781       Potential.
FT   TOPO_DOM    782    786       Extracellular (Potential).
FT   TRANSMEM    787    807       Potential.
FT   TOPO_DOM    808    840       Cytoplasmic (Potential).
FT   CARBOHYD     86     86       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     87     87       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     94     94       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    134    134       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    290    290       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    478    478       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    528    528       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   840 AA;  93497 MW;  1FCFB7EFC6BB45DB CRC64;

  Query Match             26.7%;  Score 1208.5;  DB 1;  Length 840;
  Best Local Similarity   33.8%;  Pred. No. 6.2e-79;
  Matches  291;  Conservative  151;  Mismatches  339;  Indels   79;  Gaps   18;

Qy         12 WALLHPGTGAPLCLSQQL----------------RMKGDYVLGGLFPL-GEAEEAGLRS 53
              ||       | | || ||                  : ||::| ||| | |:     |:
```

```
Db          4 WA-------AHLLLSLQLVYCWAFSCQRTESSPGFSLPGDFLLAGLFSLHGDC----LQV 52

Qy         54 RTRPSSPVCTR---FSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAM 110

Db         53 RHRPLVTSCDRPDSFNGHGYHLFQAMRFTVEEINNSSALLPNITLGYELYDVCSES-ANV 111

Qy        111 KPSLMFLAKAGSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGAS 170

Db        112 YATLRVLALQGPRHIEIQKDLRNHSSKVVAFIGPDNTDHAVTTAALLGPFLMPLVSYEAS 171

Qy        171 MELLSARETFPSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALA 230

Db        172 SVVLSAKRKFPSFLRTVPSDRHQVEVMVQLLQSFGWVWISLIGSYGDYGQLGVQALEELA 231

Qy        231 AARGICIAHEGLVPLPRADDSRLG--KVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSI 288

Db        232 VPRGICVAFKDIVPF----SARVGDPRMQSMMQHLAQARTTVVVVFSNRHLARVFFRSVV 287

Qy        289 SSRLSPKVWVASEAWLTSDLVMGLPGMAQMGTVLGFLQRGAQ---LHEFPQ-YVKTHLAL 344

Db        288 LANLTGKVWVASEDWAISTYITSVTGIQGIGTVLGVAVQQRQVPGLKEFEESYVRAVTAA 347

Qy        345 AT---DPAFCSALGEREQGLEEDVVGQRCPQCDCITLQNVS--AGLNHHQTFSVYAAVYS 399

Db        348 PSACPEGSWCS-------------TNQLCRECHTFTTRNMPTLGAFSMSAAYRVYEAVYA 394

Qy        400 VAQALHNTLQCNASGCPAQDPVKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLW 459

Db        395 VAHGLHQLLGCTSEIC-SRGPVYPWQLLQQIYKVNFLLHENTVAFDDNGDTLGYYDIIAW 453

Qy        460 VWQGSVPRLHDVGRFNGS---LRTERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHS 516

Db        454 DWNGPEWTFEIIGSASLSPVHLDINKTKIQWHGKNNQVPVSVCTTDCLAGHHRVVVGSHH 513

Qy        517 CCYDCVDCEAGSYRQNPDDIACTFCGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLL 576

Db        514 CCFECVPCEAGTFLNMSELHICQPCGTEEWAPKESTTCFPRTVEFLAWHEPISLVLIAAN 573

Qy        577 SLALGLVLAALGLFVHHRDSPLVQASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQ 636

Db        574 TLLLLLLVGTAGLFAWHFHTPVVRSAGGRLCFLMLGSLVAGSCSFYSFFGEPTVPACLLR 633

Qy        637 QPLSHLPLTGCLSTLFLQAAEIFV----ESELPL---SWADRLSGCLRGPWAWLVVLLAM 689

Db        634 QPLFSLGFAIFLSCLTIRSFQLVIIFKFSTKVPTFYRTWA-------QNHGAGLFVIVSS 686

Qy        690 LVEVALCTWYLVAFPPEVVTDWHMLPTEALVHCRTRSWVSFGLAHATNATLAFLCFLGTF 749

Db        687 TVHLLICLTWLVMWTPRPTREYQRFPHLVILECTEVNSVGFLLAFTHNILLSISTFVCSY 746

Qy        750 LVRSQPGCYNRARGLTFAMLAYFITWVSFVPLLANVQVVLRPAVQMGALLLCVLGILAAF 809

Db        747 LGKELPENYNEAKCVTFSLLLNFVSWIAFFTMASIYQGSYLPAVNVLAGLTTLSGGFSGY 806

Qy        810 HLPRCYLLMRQPGLNTPEFF 829

Db        807 FLPKCYVILCRPELNNTEHF 826


RESULT 9
TS1R1_MOUSE
ID   TS1R1_MOUSE    STANDARD;      PRT;   842 AA.
AC   Q99PG6; Q6NS58; Q923J9; Q925I5; Q99PG5;
DT   01-FEB-2005 (Rel. 46, Created)
DT   01-FEB-2005 (Rel. 46, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Taste receptor type 1 member 1 precursor (G-protein coupled receptor
DE   70).
GN   Name=Tas1r1; Synonyms=Gpr70, T1r1, Tr1;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muroidea; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE (ISOFORM 1), AND VARIANTS ASP-244; ALA-281;
RP   309-TYR--PRO-314; THR-347; TYR-418; ASN-443 AND GLU-626.
RC   STRAIN=129/J, and C57BL/6ByJ;
RX   MEDLINE=21030739; PubMed=11178737; DOI=10.1007/s003350010236;
RA   Li X., Inoue M., Reed D.R., Huque T., Puchalski R.B., Tordoff M.G.,
RA   Ninomiya Y., Beauchamp G.K., Bachmanov A.A.;
RT   "High-resolution genetic mapping of the saccharin preference locus
RT   (Sac) and the putative sweet taste receptor (T1R1) gene (Gpr70) to
RT   mouse distal Chromosome 4.";
RL   Mamm. Genome 12:13-16(2001).
RN   [2]
RP   NUCLEOTIDE SEQUENCE (ISOFORM 1), AND TISSUE SPECIFICITY.
RC   STRAIN=C57BL/6J; TISSUE=Circumvallate papillae, and Foliate papillae;
RX   MEDLINE=21219400; PubMed=11319557; DOI=10.1038/87440;
RA   Montmayeur J.-P., Liberles S.D., Matsunami H., Buck L.B.;
RT   "A candidate taste receptor gene near a sweet taste locus.";
RL   Nat. Neurosci. 4:492-498(2001).
RN   [3]
RP   NUCLEOTIDE SEQUENCE (ISOFORM 1).
RX   MEDLINE=21400448; PubMed=11509186; DOI=10.1016/S0092-8674(01)00451-2;
RA   Nelson G., Hoon M.A., Chandrashekar J., Zhang Y., Ryba N.J.P.,
RA   Zuker C.S.;
RT   "Mammalian sweet taste receptors.";
RL   Cell 106:381-390(2001).
RN   [4]
RP   NUCLEOTIDE SEQUENCE (ISOFORM 1).
RC   STRAIN=C57BL/6;
RX   PubMed=12892531; DOI=10.1290/1543-706X(2003)039<0080:TRTIAE>2.0.CO;2;
RA   Ariyasu T., Matsumoto S., Kyono F., Hanaya T., Arai S., Ikeda M.,
RA   Kurimoto M.;
RT   "Taste receptor T1R3 is an essential molecule for the cellular
RT   recognition of the disaccharide trehalose.";
RL   In Vitro Cell. Dev. Biol. Anim. 39:80-88(2003).
RN   [5]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 2).
RC   STRAIN=C57BL/6; TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [6]
RP   FUNCTION.
RX   PubMed=11894099; DOI=10.1038/nature726;
RA   Nelson G., Chandrashekar J., Hoon M.A., Feng L., Zhao G., Ryba N.J.P.,
RA   Zuker C.S.;
RT   "An amino-acid taste receptor.";
RL   Nature 416:199-202(2002).
CC   -!- FUNCTION: Putative taste receptor. TAS1R1/TAS1R3 responds to the
CC       umami taste stimulus (the taste of monosodium glutamate) and also
CC       to most of the 20 standard L-amino acids, but not to their D-
CC       enantiomers or other compounds. Sequence differences within and
CC       between species can significantly influence the selectivity and
CC       specificity of taste responses.
CC   -!- SUBUNIT: Forms heterodimers with TAS1R3.
```

```
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=Q99PG6-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q99PG6-2; Sequence=VSP_012423, VSP_012424;
CC         Note=No experimental confirmation available. May be due to an
CC         intron retention;
CC   -!- TISSUE SPECIFICITY: Expressed strongly only in fungiform papillae.
CC   -!- SIMILARITY: Belongs to the G-protein coupled receptor 3 family.
CC       TAS1R subfamily.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; AF301161; AAK07091.1; -; mRNA.
DR   EMBL; AF301162; AAK07092.1; -; mRNA.
DR   EMBL; AF337040; AAK39437.1; -; mRNA.
DR   EMBL; AY032622; AAK51603.1; -; mRNA.
DR   EMBL; BC070448; AAH70448.1; -; mRNA.
DR   HSSP; P23385; 1EWK.
DR   MGI; MGI:1927505; Tas1r1.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   GO; GO:0016021; C:integral to membrane; TAS.
DR   InterPro; IPR001828; ANF_rcpt_lig_bd.
DR   InterPro; IPR000068; Ca_sens_receptor.
DR   InterPro; IPR000337; GPCR_Mgr.
DR   InterPro; IPR011500; NCD3G_GPCR.
DR   Pfam; PF00003; 7tm_3; 1.
DR   Pfam; PF01094; ANF_receptor; 1.
DR   Pfam; PF07562; NCD3G; 1.
DR   PRINTS; PR00592; CASENSINGR.
DR   PRINTS; PR00248; GPCRMGR.
DR   PROSITE; PS00979; G_PROTEIN_RECEP_F3_1; FALSE_NEG.
DR   PROSITE; PS00980; G_PROTEIN_RECEP_F3_2; FALSE_NEG.
DR   PROSITE; PS00981; G_PROTEIN_RECEP_F3_3; FALSE_NEG.
DR   PROSITE; PS50259; G_PROTEIN_RECEP_F3_4; 1.
KW   Alternative splicing; G-protein coupled receptor; Glycoprotein;
KW   Polymorphism; Receptor; Sensory transduction; Signal; Taste;
KW   Transducer; Transmembrane.
FT   SIGNAL        1     19       Potential.
FT   CHAIN        20    842       Taste receptor type 1 member 1.
FT   TOPO_DOM     20    568       Extracellular (Potential).
FT   TRANSMEM    569    589       1 (Potential).
FT   TOPO_DOM    590    604       Cytoplasmic (Potential).
FT   TRANSMEM    605    625       2 (Potential).
FT   TOPO_DOM    626    640       Extracellular (Potential).
FT   TRANSMEM    641    661       3 (Potential).
FT   TOPO_DOM    662    681       Cytoplasmic (Potential).
FT   TRANSMEM    682    702       4 (Potential).
FT   TOPO_DOM    703    725       Extracellular (Potential).
FT   TRANSMEM    726    746       5 (Potential).
FT   TOPO_DOM    747    762       Cytoplasmic (Potential).
FT   TRANSMEM    763    783       6 (Potential).
FT   TOPO_DOM    784    789       Extracellular (Potential).
FT   TRANSMEM    790    810       7 (Potential).
FT   TOPO_DOM    811    842       Cytoplasmic (Potential).
FT   CARBOHYD     88     88       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     89     89       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     96     96       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    136    136       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    292    292       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    480    480       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    530    530       N-linked (GlcNAc. . .) (Potential).
FT   VARSPLIC    494    561       PVSVCTRDCLEGHHRLVMGSHHCCFECMPCEAGTFLNTSEL
FT                                HTCQPCGTEEWAPEGSSACFSRTVEFL -> MGQGRSPGDC
FT                                FVTRIWGVMRTQSRMGPHTSGYLRSAVLPGMPAAVLSNRCR
FT                                RGDLEGSGTDPSSQSPL (in isoform 2).
FT                                /FTId=VSP_012423.
FT   VARSPLIC    562    842       Missing (in isoform 2).
FT                                /FTId=VSP_012424.
FT   VARIANT     244    244       N -> D (in strain 129/J).
FT   VARIANT     281    281       D -> A (in strain 129/J).
FT   VARIANT     309    314       NIPNVS -> YITNVP (in strain 129/J).
FT   VARIANT     347    347       M -> T (in strain 129/J).
FT   VARIANT     418    418       H -> Y (in strain 129/J).
FT   VARIANT     443    443       K -> N (in strain 129/J).
FT   VARIANT     626    626       K -> E (in strain 129/J).
FT   CONFLICT    120    120       L -> P (in Ref. 3).
FT   CONFLICT    648    648       S -> P (in Ref. 4).
FT   CONFLICT    649    649       C -> F (in Ref. 4).
FT   CONFLICT    655    655       F -> S (in Ref. 4).
FT   CONFLICT    659    659       I -> V (in Ref. 4).
FT   CONFLICT    776    776       A -> V (in Ref. 4).
SQ   SEQUENCE   842 AA;  93426 MW;  58826C43F5DD352E CRC64;
```

Query Match 26.1%; Score 1180.5; DB 1; Length 842;
Best Local Similarity 32.9%; Pred. No. 6.8e-77;
Matches 283; Conservative 159; Mismatches 342; Indels 75; Gaps 17;

```
Qy         12 WALLHPGTGAPLCLSQQL------------------RMKGDYVLGGLFPLGEAEEAGLR 52
Db          4 WA-------AHLLLSLQLAVAYCWAFSCQRTESSPGFSLPGDFLLAGLFSL---HADCLQ 53

Qy         53 SRTRPSSPVCTR---FSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEP--V 107
Db         54 VRHRPLVTSCDRSDSFNGHGYHLFQAMRFTVEEINNSTALLPNITLGYELYDVCSESSNV 113

Qy        108 VAMKPSLMFLAKAGSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSY 167
Db        114 YA---TLRVLAQQGTGHLEMQRDLRNHSSKVVALIGPDNTDHAVTTAALLSPFLMPLVSY 170

Qy        168 GASMELLSARETFPSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFS 227
Db        171 EASSVILSGKRKFPSFLRTIPSDKYQVEVIVRLLQSFGWVWISLVGSYGDYGQLGVQALE 230

Qy        228 ALAAARGICIAHEGLVPL-PRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNY 286
Db        231 ELATPRGICVAFKNVVPLSAQAGDPRM---QRMMLRLARARTTVVVVFSNRHLDGVFFRS 287

Qy        287 SISSRLSPKVWVASEAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHL-ALA 345
Db        288 VVLANLTGKVWIASEDWAISTNIPNVSGIQGIGTVLGVAIQQRQVPGLKEFEESYVQAVM 347

Qy        346 TDPAFC---SALGEREQGLEEDVVGQRCPQCDCITLQNVS--AGLNHHQTFSVYAAVYSV 400
Db        348 GAPRTCPEGSWCG----------TNQLCRECHAFTTWNMPELGAFSMSAAYNVYEAVYAV 397

Qy        401 AQALHNTLQCNASGCPAQDPVKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWV 460
Db        398 AHGLHQLLGCTSGTC-ARGPVHPWQLLQQIYKVNFLLHKKTVAFDDKGDPLGYYDIIAWD 456

Qy        461 WQGSVPRLHDVGRFNGS---LRTERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSC 517
Db        457 WNGPEWTFEVIGSASLSPVHLDINKTKIQWHGKNNQVPVSVCTRDCLEGHHRLVMGSHHC 516

Qy        518 CYDCVDCEAGSYRQNPDDIACTFCGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLS 577
Db        517 CFECMPCEAGTFLNTSELHTCQPCGTEEWAPEGSSACFSRTVEFLGWHEPISLVLLAANT 576

Qy        578 LALGLVLAALGLFVHHRDSPLVQASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQ 637
Db        577 LLLLLLIGTAGLFAWRLHTPVVRSAGGRLCFLMLGSLVAGSCSLYSFFGKPTVPACLLRQ 636

Qy        638 PLSHLPLTGCLSTLFLQAAEIFV----ESELPL---SWADRLSGCLRGPWAWLVVLLAML 690
Db        637 PLFSLGFAIFLSCLTIRSFQLVIIFKFSTKVPTFYHTWA-------QNHGAGIFVIVSST 689

Qy        691 VEVALCTWYLVAFPPEVVTDWHMLPTEALVHCRTRSWVSFGLAHATNATLAFLCFLGTFL 750
Db        690 VHLFLCLTWLAMWTPRPTREYQRFPHLVILECTEVNSVGFLVAFAHNILLSISTFVCSYL 749
```

```
Qy        751 VRSQPGCYNRARGLTFAMLAYFITWVSFVPLLANVQVVLRPAVQMGALLLCVLGILAAFH 810
              :  |  || |: :||::| :|::|::|  : :  |    ||| : | |  : |  : :
Db        750 GKELPENYNEAKCVTFSLLLHFVSWIAFFTMSSIYQGSYLPAVNVLAGLATLSGGFSGYF 809

Qy        811 LPRCYLLMRQPGLNTPEFF 829
              ||:||::: :| ||  | |
Db        810 LPKCYVILCRPELNNTEHF 828


RESULT 10
TS1R1_HUMAN
ID   TS1R1_HUMAN    STANDARD;      PRT;   841 AA.
AC   Q7RTX1; Q8NGZ7; Q8TDJ7; Q8TDJ8; Q8TDJ9; Q8TDK0;
DT   01-FEB-2005 (Rel. 46, Created)
DT   01-FEB-2005 (Rel. 46, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Taste receptor type 1 member 1 precursor (G-protein coupled receptor
DE   70) (Gm148).
GN   Name=TAS1R1; Synonyms=GPR70, T1R1, TR1;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RA   Suwa M., Sato T., Okouchi I., Arita M., Futami K., Matsumoto S.,
RA   Tsutsumi S., Aburatani H., Asai K., Akiyama Y.;
RT   "Genome-wide discovery and analysis of human seven transmembrane helix
RT   receptor genes.";
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RG   Human chromosome 1 international sequencing consortium;
RL   Submitted (NOV-2004) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   NUCLEOTIDE SEQUENCE OF 65-841 (ISOFORMS 1; 2; 3 AND 4).
RX   MEDLINE=21888635; PubMed=11891061; DOI=10.1016/S0378-1119(01)00897-6;
RA   Makalowska I., Sood R., Faruque M.U., Hu P., Robbins C.M.,
RA   Eddings E.M., Mestre J.D., Baxevanis A.D., Carpten J.D.;
RT   "Identification of six novel genes by experimental validation of
RT   GeneMachine predicted genes.";
RL   Gene 284:203-213(2002).
RN   [4]
RP   IDENTIFICATION OF ISOFORM 1, AND FUNCTION.
RX   MEDLINE=21927605; PubMed=11917125; DOI=10.1073/pnas.072090199;
RA   Li X., Staszewski L., Xu H., Durick K., Zoller M., Adler E.;
RT   "Human receptors for sweet and umami taste.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:4692-4696(2002).
CC   -!- FUNCTION: Putative taste receptor. TAS1R1/TAS1R3 responds to the
CC       umami taste stimulus (the taste of monosodium glutamate). Sequence
CC       differences within and between species can significantly influence
CC       the selectivity and specificity of taste responses.
CC   -!- SUBUNIT: Forms heterodimers with TAS1R3.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=4;
CC       Name=1; Synonyms=Gm148 form B;
CC         IsoId=Q7RTX1-1; Sequence=Displayed;
CC       Name=2; Synonyms=Gm148 form C;
CC         IsoId=Q7RTX1-2; Sequence=VSP_012418;
CC         Note=No experimental confirmation available;
CC       Name=3; Synonyms=Gm148 form A;
CC         IsoId=Q7RTX1-3; Sequence=VSP_012419, VSP_012420;
CC         Note=No experimental confirmation available;
CC       Name=4; Synonyms=Gm148 form D;
CC         IsoId=Q7RTX1-4; Sequence=VSP_012418, VSP_012421, VSP_012422;
CC         Note=Incomplete sequence;
CC   -!- SIMILARITY: Belongs to the G-protein coupled receptor 3 family.
CC       TAS1R subfamily.
CC   -!- CAUTION: Ref.1 sequence differs from that shown due to erroneous
CC       gene model prediction.
CC   -!- DATABASE: NAME=Protein Spotlight; NOTE=Issue 55 of February 2005;
CC       WWW="http://www.expasy.org/spotlight/back_issues/sptlt055.shtml".
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; AB065618; BAC05845.1; ALT_SEQ; Genomic_DNA.
DR   EMBL; AL591866; CAI16073.1; -; Genomic_DNA.
DR   EMBL; AF387618; AAL91359.1; ALT_INIT; mRNA.
DR   EMBL; AF387619; AAL91360.1; ALT_INIT; mRNA.
DR   EMBL; AF387617; AAL91358.1; ALT_INIT; mRNA.
DR   EMBL; AF387620; AAL91361.1; ALT_INIT; mRNA.
DR   EMBL; BK000153; DAA00012.1; -; mRNA.
DR   Ensembl; ENSG00000173662; Homo sapiens.
DR   HGNC; HGNC:14448; TAS1R1.
DR   MIM; 606225; -.
DR   GO; GO:0016021; C:integral to membrane; IC.
DR   GO; GO:0046982; F:protein heterodimerization activity; NAS.
DR   GO; GO:0008527; F:taste receptor activity; IDA.
DR   GO; GO:0050917; P:perception of umami taste; IDA.
DR   InterPro; IPR001828; ANF_rcpt_lig_bd.
DR   InterPro; IPR000068; Ca_sens_receptor.
DR   InterPro; IPR000337; GPCR_Mgr.
DR   InterPro; IPR011500; NCD3G_GPCR.
DR   Pfam; PF00003; 7tm_3; 1.
DR   Pfam; PF01094; ANF_receptor; 1.
DR   Pfam; PF07562; NCD3G; 1.
DR   PRINTS; PR00592; CASENSINGR.
DR   PRINTS; PR00248; GPCRMGR.
DR   PROSITE; PS00979; G_PROTEIN_RECEP_F3_1; FALSE_NEG.
DR   PROSITE; PS00980; G_PROTEIN_RECEP_F3_2; FALSE_NEG.
DR   PROSITE; PS00981; G_PROTEIN_RECEP_F3_3; FALSE_NEG.
DR   PROSITE; PS50259; G_PROTEIN_RECEP_F3_4; 1.
KW   Alternative splicing; G-protein coupled receptor; Glycoprotein;
KW   Receptor; Sensory transduction; Signal; Taste; Transducer;
KW   Transmembrane.
FT   SIGNAL       1     20       Potential.
FT   CHAIN       21    841       Taste receptor type 1 member 1.
FT   TOPO_DOM    21    567       Extracellular (Potential).
FT   TRANSMEM   568    588       1 (Potential).
FT   TOPO_DOM   589    603       Cytoplasmic (Potential).
FT   TRANSMEM   604    624       2 (Potential).
FT   TOPO_DOM   625    639       Extracellular (Potential).
FT   TRANSMEM   640    660       3 (Potential).
FT   TOPO_DOM   661    680       Cytoplasmic (Potential).
FT   TRANSMEM   681    701       4 (Potential).
FT   TOPO_DOM   702    725       Extracellular (Potential).
FT   TRANSMEM   726    746       5 (Potential).
FT   TOPO_DOM   747    761       Cytoplasmic (Potential).
FT   TRANSMEM   762    782       6 (Potential).
FT   TOPO_DOM   783    795       Extracellular (Potential).
FT   TRANSMEM   796    816       7 (Potential).
FT   TOPO_DOM   817    841       Cytoplasmic (Potential).
FT   CARBOHYD    87     87       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    88     88       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    95     95       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   291    291       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   479    479       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   529    529       N-linked (GlcNAc. . .) (Potential).
FT   VARSPLIC   167    420       Missing (in isoform 2 and isoform 4).
FT                               /FTId=VSP_012418.
FT   VARSPLIC   340    382       AYARADKKAPRPCHKGSWCSSNQLCRECQAFMAHTMPKLKA
FT                               FS -> PMPGQTRRPLGLATRAPGAAAISSAENAKLSWHTR
FT                               CPSSKPSP (in isoform 3).
FT                               /FTId=VSP_012419.
FT   VARSPLIC   383    841       Missing (in isoform 3).
FT                               /FTId=VSP_012420.
FT   VARSPLIC   532    574       DLYRCQPCGKEEWAPEGSQTCFPRTVVFLALREHTSWVLLA
FT                               AN -> ATWVRTCQRTTTRTNVSPSACSSTSCPGSPSSPRP
FT                               ASTTAXTC (in isoform 4).
FT                               /FTId=VSP_012421.
```

```
FT   VARSPLIC    575    841       Missing (in isoform 4).
FT                                /FTId=VSP_012422.
FT   CONFLICT      1     63       MLLCTARLVGLQLLISCCWAFACHSTESSPDFTLPGDYLLA
FT                                GLFPLHSGCLQVRHRPEVTLCD -> MGESEKSSGSLAGSG
FT                                RGGWPGERASLFLLLSLRALVEVVVWGTGPMTSKVPFAPKP
FT                                QSLGGQRVSAGLCPSQLSLTGFLH (in Ref. 1).
FT   CONFLICT    167    167       I -> L (in Ref. 3; AAL91358).
FT   CONFLICT    648    648       C -> W (in Ref. 3; AAL91360).
SQ   SEQUENCE   841 AA;  93074 MW;  E800772B267D0216 CRC64;

  Query Match             25.5%;  Score 1153;  DB 1;  Length 841;
  Best Local Similarity   33.1%;  Pred. No. 6.8e-75;
  Matches  282;  Conservative  149;  Mismatches  360;  Indels   62;  Gaps   18;

Qy          6 VLGLSL-----WALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSP 60
              ::|| |     ||     | :    |    : |||:| |||||       |: | ||
Db          8 LVGLQLLISCCWAFACHSTES----SPDFTLPGDYLLAGLFPL---HSGCLQVRHRPEVT 60

Qy         61 VCTR---FSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFL 117
              :| |   |: :|     ||:: |||||| : ||| : ||| |:| ||:    :  :|  |
Db         61 LCDRSCSFNEHGYHLFQAMRLGVEEINNSTALLPNITLGYQLYDVCSDS-ANVYATLRVL 119

Qy        118 AKAGSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSAR 177
              :  |   |    :   | | ||||||| |:  |  |    | ||:| :|| || | || :
Db        120 SLPGQHHIELQGDLLHYSPTVLAVIGPDSTNRAATTAALLSPFLVPMISYAASSETLSVK 179

Qy        178 ETFPSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICI 237
              :||| ||:|:|: |:     |||:||| |:: :|| |:||: |:      |  :||||
Db        180 RQYPSFLRTIPNDKYQVETMVLLLQKFGWTWISLVGSSDDYGQLGVQALENQATGQGICI 239

Qy        238 AHEGLVPL-PRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKV 296
              | : ::|   :  | |:   | ::  : |:   ||::|:|   |   |   : : |: ||
Db        240 AFKDIMPFSAQVGDERM---QCLMRHLAQAGATVVVVFSSRQLARVFFESVVLTNLTGKV 296

Qy        297 WVASEAWLTSDLVMGLPGMAQMGTVLGF-LQRGA--QLHEFPQYVKTHLALATDP----A 349
              |||||||  |  : |:||: ::| |||  :|: |   |  | :          |   :
Db        297 WVASEAWALSRHITGVPGIQRIGMVLGVAIQKRAVPGLKAFEEAYARADKKAPRPCHKGS 356

Qy        350 FCSALGEREQGLEEDVVGQRCPQCDCI---TLQNVSAGLNHHQTFSVYAAVYSVAQALHN 406
              :||:               | | :|      |:  : |  :    :: | |||:||  ||
Db        357 WCSS-------------NQLCRECQAFMAHTMPKLKA-FSMSSAYNAYRAVYAVAHGLHQ 402

Qy        407 TLQCNASGCPAQDPVKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVP 466
              | | |||  ::  | |||||| :: : | :    : |: : :    |::  | | |
Db        403 LLGC-ASGACSRGRVYPWQLLEQIHKVHFLLHKDTVAFNDNRDPLSSYNIIAWDWNGPKW 461

Qy        467 RLHDVGRFNGS---LRTERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVD 523
                :|    |    |      ||:|| ||| | | ||  | ||  | | ||| ||::||
Db        462 TFTVLGSSTWSPVQLNINETKIQWHGKDNQVPKSVCSSDCLEGHQRVVTGFHHCCFECVP 521

Qy        524 CEAGSYRQNPDDIACTFCGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLV 583
              | ||::    |   |  ||::||:|| |  || |   |||  |    :||   :| | |:
Db        522 CGAGTFLNKSDLYRCQPCGKEEWAPEGSQTCFPRTVVFLALREHTSWVLLAANTLLLLLL 581

Qy        584 LAALGLFVHHRDSPLVQASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLP 643
              |   |||  | |:|:|:::|| |    |  |     |:  | |:|:   || :| |  |
Db        582 LGTAGLFAWHLDTPVVRSAGGRLCFLMLGSLAAGSGSLYGFFGEPTRPACLLRQALFALG 641

Qy        644 LTGCLSTLFLQAAEIFV----ESELPL---SWADRLSGCLRGPWAWLVVLLAMLVEVALC 696
              |  || | ::: :: :     :::|    :|       ::    | | |:::   :: :|
Db        642 FTIFLSCLTVRSFQLIIIFKFSTKVPTFYHAW-------VQNHGAGLFVMISSAAQLLIC 694

Qy        697 TWYLVAFPPEVVTDWHMLPTEALVHCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPG 756
              :|| : |    ::   |   :: |   : : | ||    | |:    |  ::| : |
Db        695 LTWLVVWTPLPAREYQRFPHLVMLECTETNSLGFILAFLYNGLLSISAFACSYLGKDLPE 754

Qy        757 CYNRARGLTFAMLAYFITWVSFVPLLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYL 816
              || |: :||::|  |::|::|    :        ||  | | |  :       : ||:||:
Db        755 NYNEAKCVTFSLLFNFVSWIAFFTTASVYDGKYLPAANMMAGLSSLSSGFGGYFLPKCYV 814

Qy        817 LMRQPGLNTPEFF 829
              :: :| ||: | |
Db        815 ILCRPDLNSTEHF 827
```

```
RESULT 11
Q8K4Z6_MOUSE
ID   Q8K4Z6_MOUSE PRELIMINARY;      PRT;   928 AA.
AC   Q8K4Z6;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   G protein-coupled receptor family C group 6 member A (Orphan G-protein
DE   coupled receptor isoform 1).
GN   Name=Gprc6a;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Swiss-Webster/NIH;
RA   Lorenz-Depiereux B., Dorner M., Strom T.M.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=CD-1;
RX   PubMed=15816861; DOI=10.1111/j.1471-4159.2005.03025.x;
RA   Kuang D., Yao Y., Lam J., Tsushima R.G., Hampson D.R.;
RT   "Cloning and characterization of a family C orphan G-protein coupled
RT   receptor.";
RL   J. Neurochem. 93:383-391(2005).
DR   EMBL; AY101365; AAM22280.1; -; mRNA.
DR   EMBL; AY735399; AAW68012.1; -; mRNA.
DR   Ensembl; ENSMUSG00000019905; Mus musculus.
DR   MGI; MGI:2429498; Gprc6a.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   GO; GO:0016021; C:integral to membrane; TAS.
DR   InterPro; IPR001828; ANF_receptor.
DR   InterPro; IPR000068; Ca_sens_receptor.
DR   InterPro; IPR000337; GPCR_Mgr.
DR   InterPro; IPR011500; NCD3G_GPCR.
DR   Pfam; PF00003; 7tm_3; 1.
DR   Pfam; PF01094; ANF_receptor; 1.
DR   Pfam; PF07562; NCD3G; 1.
DR   PRINTS; PR00592; CASENSINGR.
DR   PRINTS; PR00248; GPCRMGR.
DR   PROSITE; PS00980; G_PROTEIN_RECEP_F3_2; UNKNOWN_1.
DR   PROSITE; PS50259; G_PROTEIN_RECEP_F3_4; 1.
KW   Receptor.
SQ   SEQUENCE   928 AA;  104213 MW;  D670E3FEC700BD9B CRC64;

  Query Match             24.2%;  Score 1095.5;  DB 2;  Length 928;
  Best Local Similarity   31.0%;  Pred. No. 1.2e-70;
  Matches  263;  Conservative  161;  Mismatches  345;  Indels   79;  Gaps   21;

Qy         33 GDYVLGGLFPLGEAEEAGLRSRTRPSSPVCTRFSSNGLLWALAMKMAVEEINNKSDLLPG 92
              |  ::|||| : |   :       ||    |  |  :  |  |||  ::| ||| | || |
Db         34 GHIMIGGLFAIHEKMLSSDDHPRRPQIQKCAGFEISVFLQTLAMIHSIEMINN-STLLSG 92

Qy         93 LRLGYDLFDTCSEPVVAMKPSLMFLAKAG-SRDIAAY-CNYTQYQPRVLAVIGPHSSELA 150
              ::|||:::|||:|    ||   :| ||:|   ||:   : |:|: | ||| ||||    || :
Db         93 VKLGYEIYDTCTEVTAAMAATLRFLSKFNCSRETVVFQCDYSSYMPRVKAVIGAGYSETS 152

Qy        151 MVTGKFFSFFLMPQVSYGASMELLSARETFPSFFRTVPSDRVQLTAAAELLQEFGWNWVA 210
              :   :  : |||||||  :: |:|| :  |||| ||||||  |  | | |:::  ||||:
Db        153 IAVSRMLNLQLMPQVSYESTAEILSDKIRFPSFLRTVPSDFYQTKAMAHLIRQSGWNWIG 212

Qy        211 ALGSDDEYGRQGLSIFSALAAARGICIAHEGLVPLPRADDSRLGKVQDVLHQ-VNQSSVQ 269
              |: :||:|||  |: |:  ||   :||| : ::|   :|::   ::   | : : :: |
Db        213 AITTDDDYGRLALNTFAIQAAENNVCIAFKEVLPAFLSDNTIEVRINQTLEKIIAEAQVN 272

Qy        270 VVLLFASVHAAHALFNYSISSRLSPKVWVASEAWLTSDLVMGLPGMAQMGTVLGFLQRGA 329
              |:::|         ||  :|  ::| |:|:||: | |:  :: :| : ::| |:||  |
Db        273 VIVVFLRKFHVFNLFTKAIERKIS-KIWIASDNWSTATKIITIPNVKKLGKVVGFAFRRG 331
```

```
Qy        330 QLHEFPQYVKT-HL----------ALATDPAFCSALGEREQGLEEDVVGQRCPQC----- 373
Db        332 NTSSFHSFLQTLHMYPNDNNKPLHEFAMLVSACKYIKDGD----------LSQCISNYS 380

Qy        374 -DCITLQNVSAGLNH---------HQT-----FSVYAAVYSVAQALHNTLQCNASGCPAQ 418
Db        381 QATLTYDTTKTIENHLFKRNDFLWHYTEPGLIYSIQLAVFALGHAIRDL--CQARDCKKP 438

Qy        419 DPVKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRF---- 474
Db        439 NAFQPWELLAVLKNVTFTDGRNSFHFDAHGDLNTGYDVVLWKETNGLMTVTKMAEYDLQR 498

Qy        475 ------NGSLRTERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGF-HSCCYDCVDCEAG 527
Db        499 DVFITTNQETKHEFRKLK-------QILSKCSKECSPGQMKKATGSQHSCCYECVSCPEN 551

Qy        528 SYRQNPDDIACTFCGQD-EWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAA 586
Db        552 HYSNETDMDHCLLCNNETHWAPVRSTTCFEKEVEYLDWDDSLALLLIALSLLGIAFVLAI 611

Qy        587 LGLFVHHRDSPLVQASGGPLACF-GLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLT 645
Db        612 GIIFTRNLKTPVVKSSGGLVVCYVMLICHALNFASTGFFIGEPQDFACKTRQTLFGVSFT 671

Qy        646 GCLSTLFLQAAEIFVESELPLSWADRLS---GCLRGPWAWLVVLLAMLVEVALCTWYLVA 702
Db        672 LCVSCILTKSLKIL----LAFSFDPKLTMFLKCLYRPVP--IVLTCTGIQVVICTLWLVL 725

Qy        703 FPPEVVTDWHMLPTEALVHCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRAR 762
Db        726 AAPSVEEN-ISLPRVIILECEEGSALAFGTMLGYITVLAFICFVFAFKGRKLPENYNEAK 784

Qy        763 GLTFAMLAYFITWVSFVPLLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPG 822
Db        785 FLTFGMLIYFIAWITFIPVYTTTFGKYLPAVEIIVILISNYGILCCIFFPKCYIILCKQK 844

Qy        823 LNTPEFFL 830
Db        845 TNTKSAFL 852
```

RESULT 12
Q70VB1_RAT
ID  Q70VB1_RAT  PRELIMINARY;  PRT;  928 AA.
AC  Q70VB1;
DT  05-JUL-2004 (TrEMBLrel. 27, Created)
DT  05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT  05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE  Putative G protein-coupled receptor (Fragment).
GN  Name=Gprc6a; Synonyms=gprc;
OS  Rattus norvegicus (Rat).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC  Muridae; Murinae; Rattus.
OX  NCBI_TaxID=10116;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RC  STRAIN=Wistar; TISSUE=Circumvallate taste papilla;
RA  Burhenne N., Schmale H.;
RL  Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.
DR  EMBL; AJ535460; CAD59483.1; -; mRNA.
DR  RGD; 735077; Gprc6a.
DR  GO; GO:0016021; C:integral to membrane; IEA.
DR  GO; GO:0016020; C:membrane; IEA.
DR  GO; GO:0008067; F:metabotropic glutamate, GABA-B-like recepto. . .; IEA.
DR  GO; GO:0004872; F:receptor activity; IEA.
DR  GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR  InterPro; IPR001828; ANF_receptor.
DR  InterPro; IPR000068; Ca_sens_receptor.
DR  InterPro; IPR000337; GPCR_Mgr.
DR  InterPro; IPR011500; NCD3G_GPCR.
DR  InterPro; IPR004073; Vmron_receptor2.
DR  Pfam; PF00003; 7tm_3; 1.
DR  Pfam; PF01094; ANF_receptor; 1.
DR  Pfam; PF07562; NCD3G; 1.
DR  PRINTS; PR00592; CASENSINGR.
DR  PRINTS; PR00248; GPCRMGR.
DR  PRINTS; PR01535; VOMERONASL2R.
DR  PROSITE; PS00980; G_PROTEIN_RECEP_F3_2; UNKNOWN_1.
DR  PROSITE; PS50259; G_PROTEIN_RECEP_F3_4; 1.
KW  Receptor.
FT  NON_TER  1  1
SQ  SEQUENCE  928 AA;  104352 MW;  0B13BF36E6A60BC4 CRC64;

Query Match  24.0%;  Score 1086.5;  DB 2;  Length 928;
Best Local Similarity  31.7%;  Pred. No. 5.3e-70;
Matches  266;  Conservative  158;  Mismatches  353;  Indels  63;  Gaps  20;

```
Qy         33 GDYVLGGLFPLGEAEEAGLRSRTRPSSP---VCTRFSSNGLLWALAMKMAVEEINNKSDL 89
Db         34 GHIMIGGLFAI---HEKMLSSDDHPRQPQIQKCVGFEISVFLQTLAMIHSIEMINN-SSL 89

Qy         90 LPGLRLGYDLFDTCSEPVVAMKPSLMFLAKAG-SRDIAAY-CNYTQYQPRVLAVIGPHSS 147
Db         90 LSGVKLGYEIYDTCTEVTAAMAATLRFLSKFNCSRETVIFQCDYSSYVPRVKAIIGAGYS 149

Qy        148 ELAMVTGKFFSFFLMPQVSYGASMELLSARETFPSFFRTVPSDRVQLTAAAELLQEFGWN 207
Db        150 EISMAVSRMLNLQLMPQVSYESTAEILSDKIRFPSFLRTVPSDFYQTKAMAHLIRQSGWN 209

Qy        208 WVAALGSDDEYGRQGLSIFSALAAARGICIAHEGLVPLPRADDSRLGKVQDVLHQ-VNQS 266
Db        210 WVGAITTDDDYGRLALNTFAIQAAENNVCIAFKEVLPAFLSDNTIEVRINQTLEKIIAEA 269

Qy        267 SVQVVLLFASVHAAHALFNYSISSRLSPKVWVASEAWLTSDLVMGLPGMAQMGTVLGFLQ 326
Db        270 QVNVIVVFLRKFHVFNLFNKAIERKIS-KIWIASDNWSTAAKIITIPNVKKLGKVVGFTF 328

Qy        327 RGAQLHEFPQYVKTHLALATD---------PAFCSALGEREQGLEEDVVGQRCPQCDCIT 377
Db        329 RRGNMSSFHSFLQTLHMYPSDNNKPLHEFAMLFSACKHIKDGDLSQCISNYSQATWTYDT 388

Qy        378 LQNVSAGLNHHQTF-----------SVYAAVYSVAQALHNTLQCNASGCPAQDPVKPWQL 426
Db        389 TKTIETHLFKRNDFLWHYTEPGLIHSIQLAVLALGHAIRDL--CQDRDCQKPNAFQPWEL 446

Qy        427 LENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRL---------HDVG-RFNG 476
Db        447 LAVLKNVTFTDGKNSFHFDAHGDLNTGYEVVLWKETNGLMTVTKMAEYDLQHDVFITTNQ 506

Qy        477 SLRTERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGF-HSCCYDCVDCEAGSYRQNPDD 535
Db        507 ETKHEFRKLK-------QILSKCSKECIPGQMKKATGSQHSCCYECVNCPENHYSNETDM 559

Qy        536 IACTFCGQD-EWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHR 594
Db        560 DHCLVCNNETHWAPVRSTMCFEKEVEYLDWDDSLALLLIALSLLGIAFVLAVGIIFTRNL 619

Qy        595 DSPLVQASGGPLACF-GLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFL 653
Db        620 KTPVVKSSGGLVVCYVMLACHALNFASTGFFIGEPQDFTCKTRQTLFGVSFTLCVSCILT 679

Qy        654 QAAEIFVESELPLSWADRLS---GCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTD 710
Db        680 KSLKIL----LAFSFDPTLKTFLKCLYRPVP--IVLTCTGIQVVICTLWLVLAAP-TVEE 732

Qy        711 WHMLPTEALVHCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLA 770
Db        733 NTSLPRVIILECEEGSALAFGTMLGYIAVLAFICFVFAFKGRKLPENYNEAKFLTFGMLI 792

Qy        771 YFITWVSFVPLLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFL 830
Db        793 YFIAWITFIPVYATTFGKYLPAVEIIVILISNYGILCCTFFPKCYIILCKQKTNTKSVFL 852
```

```
RESULT 13
Q5YEV6_OREMO
ID   Q5YEV6_OREMO PRELIMINARY;      PRT;   940 AA.
AC   Q5YEV6;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Extracellular calcium-sensing receptor.
OS   Oreochromis mossambicus (Mozambique tilapia) (Tilapia mossambica).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidei;
OC   Cichlidae; African cichlids; Pseudocrenilabrinae; Tilapiini;
OC   Oreochromis.
OX   NCBI_TaxID=8127;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Kidney;
RX   PubMed=15456780; DOI=10.1074/jbc.M410098200;
RA   Loretz C.A., Pollina C., Hyodo S., Takei Y., Chang W., Shoback D.;
RT   "cDNA Cloning and Functional Expression of a Ca2+-sensing Receptor
RT   with Truncated C-terminal Tail from the Mozambique Tilapia
RT   (Oreochromis mossambicus).";
RL   J. Biol. Chem. 279:53288-53297(2004).
DR   EMBL; AY541693; AAT06805.1; -; mRNA.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0008067; F:metabotropic glutamate, GABA-B-like recepto. . .; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR001828; ANF_receptor.
DR   InterPro; IPR000068; Ca_sens_receptor.
DR   InterPro; IPR000337; GPCR_Mgr.
DR   InterPro; IPR011500; NCD3G_GPCR.
DR   InterPro; IPR004073; Vmron_receptor2.
DR   Pfam; PF00003; 7tm_3; 1.
DR   Pfam; PF01094; ANF_receptor; 1.
DR   Pfam; PF07562; NCD3G; 1.
DR   PRINTS; PR00592; CASENSINGR.
DR   PRINTS; PR00248; GPCRMGR.
DR   PRINTS; PR01535; VOMERONASL2R.
DR   PROSITE; PS00979; G_PROTEIN_RECEP_F3_1; 1.
DR   PROSITE; PS50259; G_PROTEIN_RECEP_F3_4; 1.
KW   Receptor.
SQ   SEQUENCE   940 AA;  105803 MW;  92FF6116FD962ED5 CRC64;

  Query Match           23.7%;  Score 1072;  DB 2;  Length 940;
  Best Local Similarity   30.3%;  Pred. No. 6.1e-69;
  Matches  263;  Conservative  153;  Mismatches  348;  Indels  104;  Gaps   24;

Qy         27 QQLRMKGDYVLGGLFPLGEAEEAGLRSR-----TRPSSPVCTRFSSNGLLWALAMKMAVE 81
              |: :| || :|||||||      |: |:      || |  | ||:  |  |  ||  |:|
Db         24 QRAQMTGDILLGGLFPL----HFGIASKDQDLAARPESTQCVRFNFRGFRWLQAMIFAIE 79

Qy         82 EINNKSDLLLPGLRLGYDLFDTCSEPVVAMKPSLMFLA--KAGSRDIAAYCNYTQYQPRVL 139
              |||| | ||| : ||| :||||:    |:: :| |:|  |  | ::  :|| | : |  :
Db         80 EINNSSTLLPNITLGYRIFDTCNTVSKALEATLSFVAQNKIDSLNLDEFCNCTDHIPSTI 139

Qy        140 AVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETFPSFFRTVPSDRVQLTAAAE 199
              ||:|   | ::         | :||:|| :|  ||| :  : || ||:|:|  | || |:
Db        140 AVVGAAGSAVSTAVANLLGLFYIPQISYASSSRLLSNKNQYKSFMRTIPTDEYQATAMAD 199

Qy        200 LLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHEGLVPLPRADDSRLGKVQDV 259
              ::: | |||| |: ||||||| |:  |       | |||     |:      |     ::: :
Db        200 IIEYFQWNWVIAVASDDEYGRPGIEKFENEMEERDICIHLNELISQYFEDH----EIKAL 255

Qy        260 LHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVASEAWLTSDLVMGLPGMAQMG 319
              : ::  |: :|:::|||       |    :   :: ::|:|||||  | ||     :  |
Db        256 VDRIENSTAKVIVVFASGPDVEPLIKEMVRRNITDRIWLASEAWAISSLVAKPEYLDVMA 315

Qy        320 TVLGFLQRGAQL---HEFPQYVKT----------------HLALATDPAFCSALGEREQ 359
                :||  :  ::    || |:|:                  :  |   | |     | |
Db        316 GTIGFALKAGRIPGFREFLQHVQPKKDSHNEFVREFWEETFNCYLEDSPRF----QESEN 371

Qy        360 GL---------EEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALHNTLQC 410
              |          |||:     |  |   |:          :::|| ||||:|||| : | |
Db        372 GSTSFRPLCSGEEDIASVETPYLDYKHLR---------ISYNVYVAVYSIAQALQDILTC 422

Qy        411 NA-------SGCPAQDPVKPWQLLENMYNLTF-HVGGLPLRFDSSGNVDMEYDLKLW--- 459
                     : |     :: ||:|: : :| : :  |  :||| : :::  | :  |
Db        423 TPGRGLFANNSCADIKKMEAWQVLKQLRHLNYTNSMGEKMRFDENSDMEANYTIINWHRS 482

Qy        460 VWQGSVPRLHDVGRFN------GSLRTERLKIRWHTSDNQKPVSRCSRQCQEGQVRR-VK 512
               |||    :|| :|         |  :| || |:    : | | ||  |: |  :  :
Db        483 TEDGSVV-FEEVGYYNMHAKRGAKLFIDRTKILWNGYSTEVPFSNCSEDCEPGTRKGIID 541

Qy        513 GFHSCCYDCVDCEAGSYRQNPDDIACTFCGQDEWSPERSTRCFRRRSRFLAWGEPAVLLL 572
              :||::| :|  | |  : |   |  |  : ||    | || :   ||:| ||    :
Db        542 SMPTCCFECTECSDGEYSNHKDASVCAKCPNNSWSNGNHTFCFLKEVEFLSWTEP--FGI 599

Qy        573 LLLLSLALGLVLAA--LGLFVHHRDSPLVQASGGPLA---CFGLVCLGLVCL-SVLLFPG 626
               | :   ||:|| |  :|:||  |::|:|:|:   |:    | |:|    | | |:| |
Db        600 ALAICAVLGVVLTAFVIGVFVRFRNTPIVKATNRELSYVLLFSLIC----CFSSSLIFIG 655

Qy        627 QPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV--ESELPLSWADRLSGCLRGPW---- 680
              :|    | :||   :    |:| : ::   : :  |:::| |         |   |
Db        656 EPQDWTCRLRQPAFGVSFVLCISCILVKTNRVLLVFEAKIPTS--------LHRKWWGLN 707

Qy        681 -AWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALVHCRTRSWVSFGLAHATNAT 739
               :|:| |   |:| :|  :|   ||    : | :     : |   | :: |
Db        708 LQFLLVFLCTFVQVMICVVWLYNAPPSSYQN-HDIDEIIFITCNEGSVMALGFLIGYTCI 766

Qy        740 LAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVPLLANVQVVLRPAVQMGALL 799
              || :||   |  |  |  :  |: :||:|| :|| |:||:|   :       ||:: |:|
Db        767 LAAICFFFAFKSRKLPENFTEAKFITFSMLIFFIVWISFIPAYFSTYGKFVSAVEVIAIL 826

Qy        800 LCVLGILAAFHLPRCYLLMRQPGLNTPE 827
                |:||     : |::: :|  || |
Db        827 ASSFGMLACIPFNKVYIILFKPSRNTIE 854


RESULT 14
CASR_HUMAN
ID   CASR_HUMAN     STANDARD;      PRT;  1078 AA.
AC   P41180; Q13912; Q16108; Q16109; Q16110; Q16379;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Extracellular calcium-sensing receptor precursor (CaSR) (Parathyroid
DE   Cell calcium-sensing receptor).
GN   Name=CASR; Synonyms=GPRC2A, PCAR1;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Pearce S.H.S., Thakker R.V.;
RL   Submitted (DEC-1994) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE, AND VARIANT GLY-990.
RC   TISSUE=Parathyroid;
RX   MEDLINE=95279439; PubMed=7759551; DOI=10.1074/jbc.270.21.12919;
RA   Garrett J.E., Capuano I.V., Hammerland L.G., Hung B.C., Brown E.M.,
RA   Hebert S.C., Nemeth E.F., Fuller F.;
RT   "Molecular cloning and functional expression of human parathyroid
RT   calcium receptor cDNAs.";
RL   J. Biol. Chem. 270:12919-12925(1995).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Kidney;
RX   MEDLINE=95408281; PubMed=7677761;
RA   Aida K., Koishi S., Tawata M., Onaya T.;
RT   "Molecular cloning of a putative Ca(2+)-sensing receptor cDNA from
RT   human kidney.";
```

RL Biochem. Biophys. Res. Commun. 214:524-529(1995).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96343808; PubMed=8756555; DOI=10.1210/en.137.9.3842;
RA Freichel M., Zink-Lorenz A., Holloschi A., Hafner M., Flockerzi V.,
RA Raue F.;
RT "Expression of a calcium-sensing receptor in a human medullary thyroid
RT carcinoma cell line and its contribution to calcitonin secretion.";
RL Endocrinology 137:3842-3848(1996).
RN [5]
RP NUCLEOTIDE SEQUENCE OF 643-908.
RX MEDLINE=96193893; PubMed=8613532;
RA Bikle D.D., Ratnam A., Mauro T., Harris J., Pillai S.;
RT "Changes in calcium responsiveness and handling during keratinocyte
RT differentiation. Potential role of the calcium receptor.";
RL J. Clin. Invest. 97:1085-1093(1996).
RN [6]
RP VARIANTS FHH GLN-185; LYS-297 AND TRP-795.
RX MEDLINE=94094324; PubMed=7916660;
RA Pollak M.R., Brown E.M., Chou Y.-H.W., Hebert S.C., Marx S.J.,
RA Steinmann B., Levi T., Seidman C.E., Seidman J.G.;
RT "Mutations in the human Ca(2+)-sensing receptor gene cause familial
RT hypocalciuric hypercalcemia and neonatal severe hyperparathyroidism.";
RL Cell 75:1297-1303(1993).
RN [7]
RP VARIANT ADH ALA-127.
RX MEDLINE=95179179; PubMed=7874174;
RA Pollak M.R., Brown E.M., Estep H.L., McLaine P.N., Kifor O., Park J.,
RA Hebert S.C., Seidman C.E., Seidman J.G.;
RT "Autosomal dominant hypocalcaemia caused by a Ca(2+)-sensing receptor
RT gene mutation.";
RL Nat. Genet. 8:303-307(1994).
RN [8]
RP VARIANTS FHH MET-62; CYS-66; MET-138; GLU-143 AND GLN-227.
RX MEDLINE=95243222; PubMed=7726161;
RA Chou Y.-H.W., Pollak M.R., Brandi M.L., Toss G., Arnqvist H.,
RA Atkinson A.B., Papapoulos S.E., Marx S., Brown E.M., Seidman J.G.,
RA Seidman C.E.;
RT "Mutations in the human Ca(2+)-sensing-receptor gene that cause
RT familial hypocalciuric hypercalcemia.";
RL Am. J. Hum. Genet. 56:1075-1079(1995).
RN [9]
RP NUCLEOTIDE SEQUENCE OF 1-61, AND VARIANT FHH ALA-39.
RX MEDLINE=95403641; PubMed=7673400; DOI=10.1210/jc.80.9.2594;
RA Aida K., Koishi S., Inoue M., Nakazato M., Tawata M., Onaya T.;
RT "Familial hypocalciuric hypercalcemia associated with mutation in the
RT human Ca(2+)-sensing receptor gene.";
RL J. Clin. Endocrinol. Metab. 80:2594-2598(1995).
RN [10]
RP VARIANTS NSHPT LEU-227 AND TYR-582.
RX MEDLINE=96292293; PubMed=8675635;
RA Pearce S.H.S., Trump D., Wooding C., Besser G.M., Chew S.L.,
RA Grant D.B., Heath D.A., Hughes I.A., Paterson C.R., Whyte M.P.,
RA Thakker R.V.;
RT "Calcium-sensing receptor mutations in familial benign hypercalcemia
RT and neonatal hyperparathyroidism.";
RL J. Clin. Invest. 96:2683-2692(1995).
RN [11]
RP VARIANTS FIH THR-116; HIS-681 AND SER-806, AND VARIANT SER-851.
RX MEDLINE=96311554; PubMed=8733126; DOI=10.1093/hmg/5.5.601;
RA Baron J., Winer K.K., Yanovski J.A., Cunningham A.W., Laue L.,
RA Zimmerman D., Cutler G.B. Jr.;
RT "Mutations in the Ca(2+)-sensing receptor gene cause autosomal
RT dominant and sporadic hypoparathyroidism.";
RL Hum. Mol. Genet. 5:601-606(1996).
RN [12]
RP VARIANT FHH ARG-174.
RX MEDLINE=97442275; PubMed=9298824;
RX DOI=10.1002/(SICI)1098-1004(1997)10:3<233::AID-HUMU9>3.3.CO;2-G;
RA Ward B.K., Stuckey B.G.A., Gutteridge D.H., Laing N.G., Pullan P.T.,
RA Ratajczak T.;
RT "A novel mutation (L174R) in the Ca2+-sensing receptor gene associated
RT with familial hypocalciuric hypercalcemia.";
RL Hum. Mutat. 10:233-235(1997).
RN [13]
RP VARIANT FIH VAL-616.
RX MEDLINE=99415602; PubMed=10487661; DOI=10.1210/jc.84.9.3036;
RA Stock J.L., Brown R.S., Baron J., Coderre J.A., Mancilla E.,
RA De Luca F., Ray K., Mericq M.V.;
RT "Autosomal dominant hypoparathyroidism associated with short stature
RT and premature osteoarthritis.";
RL J. Clin. Endocrinol. Metab. 84:3036-3040(1999).
RN [14]
RP VARIANT FHH GLU-557.
RX MEDLINE=21603857; PubMed=11762699; DOI=10.1385/ENDO:15:3:277;
RA Nakayama T., Minato M., Nakagawa M., Soma M., Tobe H., Aoi N.,
RA Kosuge K., Sato M., Ozawa Y., Kanmatsuse K., Kokubun S.;
RT "A novel mutation in Ca2+-sensing receptor gene in familial
RT hypocalciuric hypercalcemia.";
RL Endocrine 15:277-282(2001).
RN [15]
RP VARIANT ADH LYS-767, AND VARIANT GLY-990.
RX PubMed=15551332; DOI=10.1002/ajmg.a.30403;
RA Uckun-Kitapci A., Underwood L.E., Zhang J., Moats-Staats B.;
RT "A novel mutation (E767K) in the second extracellular loop of the
RT calcium sensing receptor in a family with autosomal dominant
RT hypocalcemia.";
RL Am. J. Med. Genet. 132A:125-129(2005).
CC -!- FUNCTION: Senses changes in the extracellular concentration of
CC calcium ions. The activity of this receptor is mediated by a G-
CC protein that activates a phosphatidylinositol-calcium second
CC messenger system.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P41180-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P41180-2; Sequence=VSP_002035;
CC -!- TISSUE SPECIFICITY: Found in kidney, but not in brain, lung,
CC liver, heart, skeletal muscle, or placenta.
CC -!- DISEASE: Defects in CASR are the cause of familial hypocalciuric
CC hypercalcemia, type 1 (FHH) [MIM:145980]; in which the receptor
CC has reduced activity. FHH is characterized by altered calcium
CC homeostasis. Affected individuals exhibit mild or modest
CC hypercalcemia, relative hypocalciuria, and inappropriately normal
CC PTH levels.
CC -!- DISEASE: Defects in CASR are the cause of neonatal severe primary
CC hyperparathyroidism (NSHPT) [MIM:239200]; in which the receptor
CC has reduced activity. NSHPT is a rare autosomal recessive life-
CC threatening disorder characterized by very high serum calcium
CC concentrations, skeletal demineralization, and parathyroid
CC hyperplasia. In some instances NSHPT has been demonstrated to be
CC the homozygous form of FHH.
CC -!- DISEASE: Defects in CASR are the cause of autosomal dominant
CC hypoparathyroidism (FIH) [MIM:146200]. FIH is characterized by
CC hypocalcemia and hyperphosphatemia due to inadequate secretion of
CC parathyroid hormone. Symptoms are seizures, tetany and cramps.
CC -!- SIMILARITY: Belongs to the G-protein coupled receptor 3 family.
CC ----------------------------------------------------------------------
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC ----------------------------------------------------------------------
DR EMBL; X81086; CAA56990.1; -; Genomic_DNA.
DR EMBL; U20759; AAA86503.1; -; mRNA.
DR EMBL; U20760; AAA86504.1; -; mRNA.
DR EMBL; D50855; BAA09453.1; -; mRNA.
DR EMBL; S83176; AAB46873.1; -; mRNA.
DR EMBL; S79217; AAB35262.2; -; mRNA.
DR EMBL; S68032; AAB29413.2; ALT_SEQ; Genomic_DNA.
DR EMBL; S68033; AAB29414.1; -; Genomic_DNA.
DR EMBL; S68036; AAB29415.1; -; Genomic_DNA.
DR EMBL; S81755; AAD14370.1; -; mRNA.

```
DR   HSSP; P23385; 1EWK.
DR   Ensembl; ENSG00000036828; Homo sapiens.
DR   HGNC; HGNC:1514; CASR.
DR   MIM; 601199; -.
DR   MIM; 145980; -.
DR   MIM; 239200; -.
DR   MIM; 146200; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0005886; C:plasma membrane; TAS.
DR   GO; GO:0004930; F:G-protein coupled receptor activity; TAS.
DR   GO; GO:0004435; F:phosphoinositide phospholipase C activity; TAS.
DR   GO; GO:0006874; P:calcium ion homeostasis; TAS.
DR   GO; GO:0005513; P:calcium ion sensing; TAS.
DR   GO; GO:0007635; P:chemosensory behavior; TAS.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; TAS.
DR   GO; GO:0009653; P:morphogenesis; TAS.
DR   GO; GO:0001503; P:ossification; TAS.
DR   InterPro; IPR001828; ANF_rcpt_lig_bd.
DR   InterPro; IPR000068; Ca_sens_receptor.
DR   InterPro; IPR000337; GPCR_Mgr.
DR   InterPro; IPR011500; NCD3G_GPCR.

  Query Match             23.6%;  Score 1068.5;  DB 1;  Length 1078;
  Best Local Similarity   30.3%;  Pred. No. 1.3e-68;
  Matches  269;  Conservative  156;  Mismatches  366;  Indels   97;  Gaps   26;

Qy         12 WALL----HPGTGAPLCLSQQLRMKGDYVLGGLFPLG---EAEEAGLRSRTRPSSPVCTR 64
Db          8 WVLLALTWHTSAYGP---DQRAQKKGDIILGGLFPIHFGVAAKDQDLKS--RPESVECIR 62

Qy         65 FSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLA--KAGS 122
Db         63 YNFRGFRWLQAMIFAIEEINSSPALLPNLTLGYRIFDTCNTVSKALEATLSFVAQNKIDS 122

Qy        123 RDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETFPS 182
Db        123 LNLDEFCNCSEHIPSTIAVVGATGSGVSTAVANLLGLFYIPQVSYASSSRLLSNKNQFKS 182

Qy        183 FFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHEGL 242
Db        183 FLRTIPNDEHQATAMADIIEYFRWNWVGTIAADDDYGRPGIEKFREEAEERDICIDFSEL 242

Qy        243 VPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVASEA 302
Db        243 I----SQYSDEEEIQHVVEVIQNSTAKVIVVFSSGPDLEPLIKEIVRRNITGKIWLASEA 298

Qy        303 WLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYV-KTHLALATDPAFCSALGERE--- 358
Db        299 WASSSLIAMPQYFHVVGGTIGFALKAGQIPGFREFLKKVHPRKSVHNGFAKEFWEETFNC 358

Qy        359 ----------------QGLEEDVVGQRCPQCD------CITLQNVSA------GLNHHQ- 389
Db        359 HLQEGAKGPLPVDTFLRGHEES--GDRFSNSSTAFRPLCTGDENISSVETPYIDYTHLRI 416

Qy        390 TFSVYAAVYSVAQALHNTLQC--------NASGCPAQDPVKPWQLLENMYNLTF-HVGGL 440
Db        417 SYNVYLAVYSIAHALQDIYTCLPGRGLFTNGS-CADIKKVEAWQVLKHLRHLNFTNNMGE 475

Qy        441 PLRFDSSGNVDMEYDLKLW---VWQGSVPRLHDVGRFN-GSLRTERL-----KIRWHTSD 491
Db        476 QVTFDECGDLVGNYSIINWHLSPEDGSIV-FKEVGYYNVYAKKGERLFINEEKILWSGFS 534

Qy        492 NQKPVSRCSRQCQEGQVRR-VKGFHSCCYDCVDCEAGSYRQNPDDIACTFCGQDEWSPER 550
Db        535 REVPFSNCSRDCLAGTRKGIIEGEPTCCFECVECPDGEYSDETDASACNKCPDDFWSNEN 594

Qy        551 STRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQASGGPLA--- 607
Db        595 HTSCIAKEIEFLSWTEPFGIALTLFAVLGIFLTAFVLGVFIKFRNTPIVKATNRELSYLL 654

Qy        608 CFGLVCLGLVCL-SVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV--ESEL 664
Db        655 LFSLLC----CFSSSLFFIGEPQDWTCRLRQPAFGISFVLCISCILVKTNRVLLVFEAKI 710

Qy        665 PLS-----WADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEAL 719
Db        711 PTSFHRKWWGLNLQ--------FLLVFLCTFMQIVICVIWLYTAPPSSYRNQELEDEIIF 762

Qy        720 VHCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFV 779
Db        763 ITCHEGSLMALGFLIGYTCLLAAICFFFAFKSRKLPENFNEAKFITFSMLIFFIVWISFI 822

Qy        780 PLLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPE 827
Db        823 PAYASTYGKFVSAVEVIAILAASFGLLACIFFNKIYIILFKPSRNTIE 870


RESULT 15
Q4PJ19_HUMAN
ID   Q4PJ19_HUMAN PRELIMINARY;      PRT;  1078 AA.
AC   Q4PJ19;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Calcium-sensing receptor (Hypocalciuric hypercalcemia 1, severe
DE   neonatal hyperparathyroidism).
GN   Name=CASR;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Rieder M.J., Johanson E.J., da Ponte S.H., Hastings N.C., Ahearn M.O.,
RA   Bertucci C.B., Wong M.W., Yi Q., Nickerson D.A.;
RL   Submitted (JUN-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; DQ088967; AAY68221.1; -; Genomic_DNA.
KW   Receptor.
SQ   SEQUENCE   1078 AA;  120674 MW;  620C57DB1A83E1DB CRC64;

  Query Match             23.6%;  Score 1068.5;  DB 2;  Length 1078;
  Best Local Similarity   30.3%;  Pred. No. 1.3e-68;
  Matches  269;  Conservative  156;  Mismatches  366;  Indels   97;  Gaps   26;

Qy         12 WALL----HPGTGAPLCLSQQLRMKGDYVLGGLFPLG---EAEEAGLRSRTRPSSPVCTR 64
Db          8 WVLLALTWHTSAYGP---DQRAQKKGDIILGGLFPIHFGVAAKDQDLKS--RPESVECIR 62

Qy         65 FSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLA--KAGS 122
Db         63 YNFRGFRWLQAMIFAIEEINSSPALLPNLTLGYRIFDTCNTVSKALEATLSFVAQNKIDS 122

Qy        123 RDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETFPS 182
Db        123 LNLDEFCNCSEHIPSTIAVVGATGSGVSTAVANLLGLFYIPQVSYASSSRLLSNKNQFKS 182

Qy        183 FFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHEGL 242
Db        183 FLRTIPNDEHQATAMADIIEYFRWNWVGTIAADDDYGRPGIEKFREEAEERDICIDFSEL 242

Qy        243 VPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVASEA 302
Db        243 I----SQYSDEEEIQHVVEVIQNSTAKVIVVFSSGPDLEPLIKEIVRRNITGKIWLASEA 298

Qy        303 WLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYV-KTHLALATDPAFCSALGERE--- 358
Db        299 WASSSLIAMPQYFHVVGGTIGFALKAGQIPGFREFLKKVHPRKSVHNGFAKEFWEETFNC 358

Qy        359 ----------------QGLEEDVVGQRCPQCD------CITLQNVSA------GLNHHQ- 389
Db        359 HLQEGAKGPLPVDTFLRGHEES--GDRFSNSSTAFRPLCTGDENISSVETPYIDYTHLRI 416

Qy        390 TFSVYAAVYSVAQALHNTLQC--------NASGCPAQDPVKPWQLLENMYNLTF-HVGGL 440
Db        417 SYNVYLAVYSIAHALQDIYTCLPGRGLFTNGS-CADIKKVEAWQVLKHLRHLNFTNNMGE 475
```

```
Qy        441 PLRFDSSGNVDMEYDLKLW---VWQGSVPRLHDVGRFN-GSLRTERL-----KIRWHTSD 491
Db        476 QVTFDECGDLVGNYSIINWHLSPEDGSIV-FKEVGYYNVYAKKGERLFINEEKILWSGFS 534

Qy        492 NQKPVSRCSRQCQEGQVRR-VKGFHSCCYDCVDCEAGSYRQNPDDIACTFCGQDEWSPER 550
Db        535 REVPFSNCSRDCLAGTRKGIIEGEPTCCFECVECPDGEYSDETDASACNKCPDDFWSNEN 594

Qy        551 STRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQASGGPLA--- 607
Db        595 HTSCIAKEIEFLSWTEPFGIALTLFAVLGIFLTAFVLGVFIKFRNTPIVKATNRELSYLL 654

Qy        608 CFGLVCLGLVCL-SVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV--ESEL 664
Db        655 LFSLLC----CFSSSLFFIGEPQDWTCRLRQPAFGISFVLCISCILVKTNRVLLVFEAKI 710

Qy        665 PLS-----WADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEAL 719
Db        711 PTSFHRKWWGLNLQ--------FLLVFLCTFMQIVICVIWLYTAPPSSYRNQELEDEIIF 762

Qy        720 VHCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFV 779
Db        763 ITCHEGSLMALGFLIGYTCLLAAICFFFAFKSRKLPENFNEAKFITFSMLIFFIVWISFI 822

Qy        780 PLLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPE 827
Db        823 PAYASTYGKFVSAVEVIAILAASFGLLACIFFNKIYIILFKPSRNTIE 870
```

Search completed: December 16, 2005, 18:10:37
Job time : 242 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2005, 17:58:51 ; Search time 44 Seconds
(without alignments)
1863.106 Million cell updates/sec

Title: US-10-679-102-15
Perfect score: 4524
Sequence: 1 MLGPAVLGLSLWALLHPGTG.....GPGDAQGQNDGNTGNQGKHE 852

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 3229 | 71.4 | 858 | 2 | JC7683 | taste receptor T1R |
| 2 | 1068.5 | 23.6 | 1078 | 2 | A56715 | calcium receptor ( |
| 3 | 1060 | 23.4 | 1085 | 2 | S40476 | Ca(2+)-sensing rec |
| 4 | 1057.5 | 23.4 | 1088 | 2 | B56715 | calcium receptor ( |
| 5 | 1054 | 23.3 | 1079 | 2 | I59362 | calcium/polyvalent |
| 6 | 670.5 | 14.8 | 872 | 2 | JH0561 | metabotropic gluta |
| 7 | 631.5 | 14.0 | 871 | 2 | A46742 | metabotropic gluta |
| 8 | 619.5 | 13.7 | 912 | 2 | JH0563 | metabotropic gluta |
| 9 | 612.5 | 13.5 | 1218 | 2 | S71376 | glutamate receptor |
| 10 | 604.5 | 13.4 | 915 | 2 | A49874 | metabotropic gluta |
| 11 | 597.5 | 13.2 | 879 | 2 | JH0562 | metabotropic gluta |
| 12 | 593.5 | 13.1 | 908 | 2 | I49142 | metabotropic gluta |
| 13 | 592.5 | 13.1 | 879 | 2 | JC7160 | metabotropic gluta |
| 14 | 584 | 12.9 | 1171 | 2 | A42916 | metabotropic gluta |
| 15 | 580.5 | 12.8 | 1180 | 2 | JC2132 | metabotropic gluta |
| 16 | 580.5 | 12.8 | 1212 | 2 | JC2131 | metabotropic gluta |
| 17 | 574.5 | 12.7 | 1199 | 2 | A41939 | G protein-coupled |
| 18 | 474 | 10.5 | 999 | 2 | T27628 | hypothetical prote |
| 19 | 436.5 | 9.6 | 1267 | 2 | T21340 | hypothetical prote |
| 20 | 323 | 7.1 | 551 | 2 | T30806 | metabotropic gluta |
| 21 | 240 | 5.3 | 962 | 2 | D86186 | hypothetical prote |
| 22 | 238 | 5.3 | 976 | 2 | T51137 | ionotropic glutama |
| 23 | 237.5 | 5.2 | 925 | 2 | T06128 | hypothetical prote |
| 24 | 236.5 | 5.2 | 941 | 2 | T51135 | ligand-gated chann |
| 25 | 221.5 | 4.9 | 953 | 2 | E84732 | probable ligand-ga |
| 26 | 220.5 | 4.9 | 1039 | 2 | T45779 | probable glutamate |
| 27 | 219 | 4.8 | 951 | 2 | T51132 | probable glutamate |
| 28 | 216.5 | 4.8 | 925 | 2 | T51133 | ligand gated chann |
| 29 | 215.5 | 4.8 | 933 | 2 | C96495 | probable ligand-ga |
| 30 | 213 | 4.7 | 950 | 2 | T51134 | ionotropic glutama |
| 31 | 213 | 4.7 | 975 | 2 | A84550 | probable ligand-ga |
| 32 | 211 | 4.7 | 923 | 2 | F84732 | probable ligand-ga |
| 33 | 210 | 4.6 | 912 | 2 | T51131 | ligand gated chann |
| 34 | 210 | 4.6 | 921 | 2 | T51136 | ionotropic glutama |
| 35 | 208.5 | 4.6 | 1099 | 2 | T16283 | hypothetical prote |
| 36 | 205 | 4.5 | 960 | 2 | JE0356 | gamma-aminobutyric |
| 37 | 175 | 3.9 | 1103 | 2 | JC5581 | guanylate cyclase |
| 38 | 159.5 | 3.5 | 1108 | 2 | I59385 | guanylate cyclase |
| 39 | 151.5 | 3.3 | 1108 | 2 | B55915 | guanylate cyclase |
| 40 | 140 | 3.1 | 965 | 2 | I51244 | N-methyl-D-asparta |
| 41 | 139.5 | 3.1 | 920 | 2 | B84640 | probable ligand-ga |
| 42 | 135.5 | 3.0 | 1323 | 2 | I78557 | N-methyl-D-asparta |
| 43 | 133.5 | 3.0 | 1356 | 1 | C45219 | N-methyl-D-asparta |
| 44 | 130.5 | 2.9 | 464 | 2 | S75362 | hypothetical prote |
| 45 | 130 | 2.9 | 938 | 2 | A46612 | N-methyl-D-asparta |

ALIGNMENTS

RESULT 1
JC7683
taste receptor T1R3 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: JC7683
R;Kitagawa, M.; Kusakabe, Y.; Miura, H.; Ninomiya, Y.; Hino, A.
Biochem. Biophys. Res. Commun. 283, 236-242, 2001
A;Title: Molecular genetic identification of a candidate receptor gene for sweet taste.
A;Reference number: JC7683; MUID:21222875; PMID:11322794
A;Contents: tongue
A;Accession: JC7683
A;Molecule type: mRNA
A;Residues: 1-858 <KIT>
A;Cross-references: UNIPROT:Q91VA4; UNIPARC:UPI00000230BB; DDBJ:AB049994
C;Comment: This protein, a seven-transmembrane receptor, serves as a mediator of the sig
C;Genetics:
A;Gene: t1r3
A;Map position: 4
C;Superfamily: metabotropic glutamate receptor 4
C;Keywords: transmembrane protein

Query Match 71.4%; Score 3229; DB 2; Length 858;
Best Local Similarity 72.5%; Pred. No. 8.3e-244;
Matches 621; Conservative 67; Mismatches 147; Indels 22; Gaps 4;

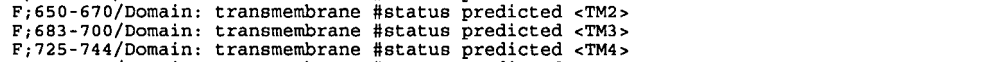
```
Qy          5 AVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSPVCTR 64
              |::|||| | |  | || |||||| : :|||:||||||||  ||| |  ||:|:|  | |
Db          5 AIMGLSLAAFLELGMGASLCLSQQFKAQGDYILGGLFPLGSTEEATLNQRTQPNSIPCNR 64

Qy         65 FSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKAGSRD 124
              ||  ||  |:|||||||||| | |||||||||||||||||||| || |||||||  ||:
Db         65 FSPLGLFLAMAMKMAVEEINNGSALLPGLRLGYDLFDTCSEPVVTMKSSLMFLAKVGSQS 124

Qy        125 IAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETFPSFF 184
              |||||||||||||||||||||||::|||||||||||||| |||: || ||||||||
Db        125 IAAYCNYTQYQPRVLAVIGPHSSELALITGKFFSFFLMPQVSYSASMDRLSDRETFPSFF 184

Qy        185 RTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHEGLVP 244
              |||||||||| |   ||| | ||||||||||:|||:||||||:|| ||||||||||||
Db        185 RTVPSDRVQLQAVVTLLQNFSWNWVAALGSDDDYGREGLSIFSSLANARGICIAHEGLVP 244

Qy        245 LPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVASEAWL 304
                    :|||| ||| ||||| ||||:||||  | ::||:|||   ||||||||||:||
Db        245 QHDTSGQQLGKVLDVLRQVNQSKVQVVVLFASARAVYSLFSYSIHHGLSPKVWVASESWL 304

Qy        305 TSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTHLALATDPAFCSALGEREQGLEED 364
              |||||| || :|::|||||||||| | ||  ||:|||||| |||||::|   |  |||
Db        305 TSDLVMTLPNIARVGTVLGFLQRGALLPEFSHYVETHLALAADPAFCASL-NAELDLEEH 363

Qy        365 VVGQRCPQCDCITLQNVSAGL--------NHHQTFSVYAAVYSVAQALHNTLQCNASGCP 416
              |:||||:|| | |||:|:||         ||| |: ||||||||||||||||| | |
```

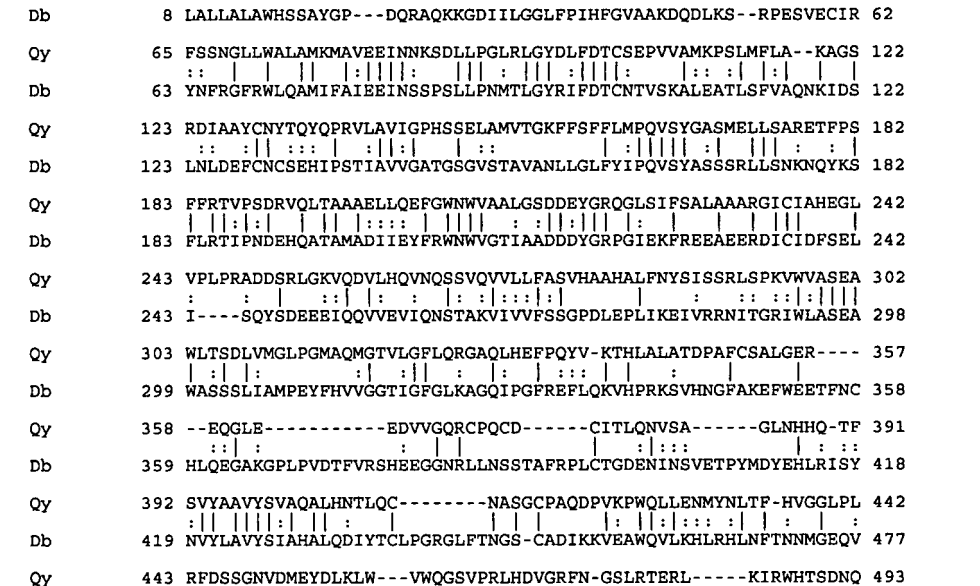
```
Db        364 VMGQRCPRCDDIMLQNLSSGLLQNLSAGQLHHQIFATYAAVYSVAQALHNTLQCNVSHCH 423

Qy        417 AQDPVKPWQLLENMYNLTFHVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNG 476
              : | ||||||||||::||   | |:||: ||||||||||:||||   | || || |||
Db        424 VSEHVLPWQLLENMYNMSFHARDLTLQFDAEGNVDMEYDLKMWVWQSPTPVLHTVGTFNG 483

Qy        477 SLRTERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVDCEAGSYRQNPDDI 536
              :|: :: |: |    || |||:||||::|||||||||||||||||||:|||||::|||
Db        484 TLQLQQSKMYW--PGNQVPVSQCSRQCKDGQVRRVKGFHSCCYDCVDCKAGSYRKHPDDF 541

Qy        537 ACTFCGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDS 596
              || | ||:||||:|| |  || :||||||| || ||||| | ||| |||||| ||| ||
Db        542 TCTPCNQDQWSPEKSTACLPRRPKFLAWGEPVVLSLLLLCLVLGLALAALGLSVHHWDS 601

Qy        597 PLVQASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAA 656
              ||||||||   ||||:|||| |||||||||:|| | |||||||:|||||||||||||||
Db        602 PLVQASGGSQFCFGLICLGLFCLSVLLFPGRPSSASCLAQQPMAHLPLTGCLSTLFLQAA 661

Qy        657 EIFVESELPLSWADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPT 716
              | |||||||||||: |   ||| |||||||||   || ||| |||:|||||||||| :|||
Db        662 ETFVESELPLSWANWLCSYLRGLWAWLVVLLATFVEAALCAWYLIAFPPEVVTDWSVLPT 721

Qy        717 EALVHCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWV 776
              | | ||  ||||| || | ||| ||||||||||||:|||| ||||||||||||||||||
Db        722 EVLEHCHVRSWVSLGLVHITNAMLAFLCFLGTFLVQSQPGRYNRARGLTFAMLAYFITWV 781

Qy        777 SFVPLLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFFLG----- 831
              ||||||||||  :|||||||:|:| |||| ||||:||:|:  | ||| |||||
Db        782 SFVPLLANVQVAYQPAVQMGAILVCALGILVTFHLPKCYVLLWLPKLNTQEFFLGRNAKK 841

Qy        832 ------GGPGDAQGQND 842
                    ||   ||| |:
Db        842 AADENSGGGEAAQGHNE 858


RESULT 2
A56715
calcium receptor (clone phPCaR-4.0) - human
C;Species: Homo sapiens (man)
C;Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 01-Dec-2000
C;Accession: A56715; S49341; A49419; B49419; C49419
R;Garrett, J.E.; Capuano, I.V.; Hammerland, L.G.; Hung, B.C.P.; Brown, E.M.; Hebert, S.C
J. Biol. Chem. 270, 12919-12925, 1995
A;Title: Molecular cloning and functional expression of human parathyroid calcium recept
A;Reference number: A56715; MUID:95279439; PMID:7759551
A;Accession: A56715
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1078 <GAR>
A;Cross-references: UNIPARC:UPI000004A99A; GB:U20759; NID:g683744; PIDN:AAA86503.1; PID:
R;Pearce, S.H.S.; Thakker, R.V.
submitted to the EMBL Data Library, August 1994
A;Reference number: S49341
A;Accession: S49341
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-180,'Q',182-989,'R',991-1078 <PEA>
A;Cross-references: UNIPARC:UPI000017C0A8; EMBL:X81086
R;Pollak, M.R.; Brown, E.M.; Chou, Y.H.; Hebert, S.C.; Marx, S.J.; Steinmann, B.; Levi,
Cell 75, 1297-1303, 1993
A;Title: Mutations in the human Ca(2+)-sensing receptor gene cause familial hypocalciuri
A;Reference number: A49419; MUID:94094324; PMID:7916660
A;Accession: A49419
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 178-192 <POL>
A;Cross-references: UNIPARC:UPI000017C0A9
A;Experimental source: family N
A;Note: sequence inconsistent with nucleotide translation
A;Note: sequence modified after extraction from NCBI backbone
A;Note: 186-Arg mutation is associated with familial hypocalciuric hypercalcemia and nec
A;Note: sequence extracted from NCBI backbone (NCBIN:142453)
A;Accession: B49419
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 289-303 <PO2>
A;Cross-references: UNIPARC:UPI000017C0AA
A;Experimental source: family E
A;Note: sequence modified after extraction from NCBI backbone
A;Note: 298-Lys mutation is associated with familial hypocalciuric hypercalcemia and neo
A;Note: sequence extracted from NCBI backbone (NCBIN:142455)
A;Accession: C49419
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 788-802 <PO3>
A;Cross-references: UNIPARC:UPI000017C0AB
A;Experimental source: family J
A;Note: sequence modified after extraction from NCBI backbone
A;Note: 796-Trp mutation is associated with familial hypocalciuric hypercalcemia and neo
A;Note: sequence extracted from NCBI backbone (NCBIN:142457)
C;Keywords: glycoprotein; receptor; transmembrane protein

  Query Match            23.6%;  Score 1068.5;  DB 2;  Length 1078;
  Best Local Similarity  30.3%;  Pred. No. 5.2e-75;
  Matches  269;  Conservative  156;  Mismatches  366;  Indels   97;  Gaps   26;

Qy         12 WALL----HPGTGAPLCLSQQLRMKGDYVLGGLFPLG---EAEEAGLRSRTRPSSPVCTR 64
              | ||    |     |    |: : ||| :||||||:     |::  |:|  || |  | |
Db          8 WVLLALTWHTSAYGP---DQRAQKKGDIILGGLFPIHFGVAAKDQDLKS--RPESVECIR 62

Qy         65 FSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLA--KAGS 122
              ::  |  |  ||  |:||||:    ||| | ||| :|||:    |:: :| |:|  |  |
Db         63 YNFRGFRWLQAMIFAIEEINSSPALLPNLTLGYRIFDTCNTVSKALEATLSFVAQNKIDS 122

Qy        123 RDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETFPS 182
              ::  :|| ::: |  :||:|   | ::          | :||||| :|  ||| :  | |
Db        123 LNLDEFCNCSEHIPSTIAVVGATGSGVSTAVANLLGLFYIPQVSYASSSRLLSNKNQFKS 182

Qy        183 FFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHEGL 242
              | ||:|:|  | || |:::: | ||||  : :||:||| |:  |   |  | |||    |
Db        183 FLRTIPNDEHQATAMADIIEYFRWNWVGTIAADDDYGRPGIEKFREEAEERDICIDFSEL 242

Qy        243 VPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVASEA 302
              :   :  |   ::| |:  :  |: :|:::|:|       |    :   :: |:|:||||
Db        243 I----SQYSDEEEIQHVVEVIQNSTAKVIVVFSSGPDLEPLIKEIVRRNITGKIWLASEA 298

Qy        303 WLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYV-KTHLALATDPAFCSALGERE--- 358
              | :| |:          :|  :||  :  |:  | ::: | |   :    |     |
Db        299 WASSSLIAMPQYFHVVGGTIGFALKAGQIPGFREFLKKVHPRKSVHNGFAKEFWEETFNC 358

Qy        359 ----------------QGLEEDVVGQRCPQCD------CITLQNVSA------GLNHHQ- 389
                           :| ||   | |             |   :|:|:          | :
Db        359 HLQEGAKGPLPVDTFLRGHEES--GDRFSNSSTAFRPLCTGDENISSVETPYIDYTHLRI 416

Qy        390 TFSVYAAVYSVAQALHNTLQC--------NASGCPAQDPVKPWQLLENMYNLTF-HVGGL 440
              :::|| ||||:| || :   |         | | |      |: ||:|::: :| | :  |
Db        417 SYNVYLAVYSIAHALQDIYTCLPGRGLFTNGS-CADIKKVEAWQVLKHLRHLNFTNNMGE 475

Qy        441 PLRFDSSGNVDMEYDLKLW---VWQGSVPRLHDVGRFN-GSLRTERL-----KIRWHTSD 491
              : ||  |::   | :  |       ||:    :|| :|  : : |||      || |
Db        476 QVTFDECGDLVGNYSIINWHLSPEDGSIV-FKEVGYYNVYAKKGERLFINEEKILWSGFS 534

Qy        492 NQKPVSRCSRQCQEGQVRR-VKGFHSCCYDCVDCEAGSYRQNPDDIACTFCGQDEWSPER 550
              : | | ||| |  |  :  ::|  :||::||:|  | |     | ||  |  | || |
Db        535 REVPFSNCSRDCLAGTRKGIIEGEPTCCFECVECPDGEYSDETDASACNKCPDDFWSNEN 594

Qy        551 STRCFRRRSRFLAWGEPAVLLLLLLSLALGLVLAALGLFVHHRDSPLVQASGGPLA--- 607
              | |  :   ||:| ||  : | |   | : |    ||:|:  |::|:|:|:   |:
Db        595 HTSCIAKEIEFLSWTEPFGIALTLFAVLGIFLTAFVLGVFIKFRNTPIVKATNRELSYLL 654

Qy        608 CFGLVCLGLVCL-SVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFV--ESEL 664
              | |:|   | | | | |:|   | :||   :    |:| : ::   : :  |:::
Db        655 LFSLLC----CFSSSLFFIGEPQDWTCRLRQPAFGISFVLCISCILVKTNRVLLVFEAKI 710

Qy        665 PLS-----WADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEAL 719
              | |     |   |         :|:| |   ::: :|  :|   ||    :  :
```

```
Db        711 PTSFHRKWWGLNLQ--------FLLVFLCTFMQIVICVIWLYTAPPSSYRNQELEDEIIF 762

Qy        720 VHCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFV 779
              : |   | :: |         || :||   |  |  | :| |: :||:|| :|| |:||:
Db        763 ITCHEGSLMALGFLIGYTCLLAAICFFFAFKSRKLPENFNEAKFITFSMLIFFIVWISFI 822

Qy        780 PLLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPE 827
              |  |:       ||:: |:|    |:||     : |::: :|  || |
Db        823 PAYASTYGKFVSAVEVIAILAASFGLLACIFFNKIYIILFKPSRNTIE 870


RESULT 3
S40476
Ca(2+)-sensing receptor - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C;Accession: S40476
R;Brown, E.M.; Gamba, G.; Riccardi, D.; Lombardi, M.; Butters, R.; Kifor, O.; Sun, A.; H
Nature 366, 575-580, 1993
A;Title: Cloning and characterization of an extracellular Ca(2+)-sensing receptor from b
A;Reference number: S40476; MUID:94077182; PMID:8255296
A;Accession: S40476
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1085 <BRO>
A;Cross-references: UNIPROT:P35384; UNIPARC:UPI000012700C; GB:S67307; NID:g453108; PIDN:

  Query Match             23.4%;  Score 1060;  DB 2;  Length 1085;
  Best Local Similarity   30.0%;  Pred. No. 2.4e-74;
  Matches  269;  Conservative  155;  Mismatches  362;  Indels  112;  Gaps   26;

Qy          6 VLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPL--GEA-EEAGLRSRTRPSSPVC 62
              :|  | |      |        |: : ||| :||||||:  | | ::  |:|  || |  |
Db         10 LLAFSTWCTSAYGP------DQRAQKKGDIILGGLFPIHFGVAVKDQDLKS--RPESVEC 61

Qy         63 TRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLA--KA 120
              |::  |  |  ||  |:||||:   ||| : ||| :||||:    |:: :| |:|  |
Db         62 IRYNFRGFRWLQAMIFAIEEINSSPALLPNMTLGYRIFDTCNTVSKALEATLSFVAQNKI 121

Qy        121 GSRDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETF 180
              | ::  :|| ::: |  :||:|   | ::          | :|||||  :|  ||| :  |
Db        122 DSLNLDEFCNCSEHIPSTIAVVGATGSGISTAVANLLGLFYIPQVSYASSSRLLSNKNQF 181

Qy        181 PSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHE 240
              || ||:|:|  | || |::::  | ||||  : :||:||| |:  |   |  | |||
Db        182 KSFLRTIPNDEHQATAMADIIEYFRWNWVGTIAADDDYGRPGIEKFREEAEERDICIDFS 241

Qy        241 GLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVAS 300
              |:    :  |   |:| |:  :  |: :|:::|:|       |    :   :: ::|:||
Db        242 ELI----SQYSDEEKIQQVVEVIQNSTAKVIVVFSSGPDLEPLIKEIVRRNITGRIWLAS 297

Qy        301 EAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYV-KTHLALATDPAFCSAL----- 354
              ||| :| |:        :|  :||  :  |:  | ::: | |   :     |
Db        298 EAWASSSLIAMPEYFHVVGGTIGFGLKAGQIPGFREFLQKVHPRKSVHNGFAKEFWEETF 357

Qy        355 ------------------GEREQGL---------------EEDVVGQRCPQCDCITLQN 380
                               |  | |                 ||::     |  |   |:
Db        358 NCHLQEGAKGPLPVDTFLRGHEEGGARLSNSPTAFRPLCTGEENISSVETPYMDYTHLR- 416

Qy        381 VSAGLNHHQTFSVYAAVYSVAQALHNTLQC--------NASGCPAQDPVKPWQLLENMYN 432
                  :::|| ||||:| || :   |        | | |     |: ||:|::: :
Db        417 --------ISYNVYLAVYSIAHALQDIYTCIPGRGLFTNGS-CADIKKVEAWQVLKHLRH 467

Qy        433 LTFHVG-GLPLRFDSSGNVDMEYDLKLW---VWQGSVPRLHDVGRFN-GSLRTERL---- 483
              | |    |  : || |::   | :  |      ||:    :|| :|  : : |||
Db        468 LNFTSNMGEQVTFDECGDLAGNYSIINWHLSPEDGSIV-FKEVGYYNVYAKKGERLFIND 526

Qy        484 -KIRWHTSDNQKPVSRCSRQCQEGQVRR-VKGFHSCCYDCVDCEAGSYRQNPDDIACTFC 541
               || |     : | | ||| |  |  :  ::|  :||::||:|  | |     |  ||  |
Db        527 EKILWSGFSREVPFSNCSRDCLAGTRKGIIEGEPTCCFECVECPDGEYSDETDASACDKC 586

Qy        542 GQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQA 601
                | || |  | |  :   ||:| ||  : | |   | : |    ||:|:  |::|:|:|

Db        587 PDDFWSNENHTSCIAKEIEFLSWTEPFGIALTLFAVLGIFLTAFVLGVFIKFRNTPIVKA 646

Qy        602 SGGPLA---CFGLVCLGLVCL-SVLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAE 657
              :   |:    | |:|    |  | | | |:|    |  :||   :    |:| : ::
Db        647 TNRELSYLLLFSLLC----CFSSSLFFIGEPQDWTCRLRQPAFGISFVLCISCILVKTNR 702

Qy        658 IFV--ESELPLS-----WADRLSGCLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTD 710
              : :  |:::| |     |  |          :|:| |   ::: :|  :|   ||    :
Db        703 VLLVFEAKIPTSFHRKWWGLNLQ--------FLLVFLCTFMQIVICAIWLNTAPPSSYRN 754

Qy        711 WHMLPTEAL-VHCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAML 769
               | |  | : : |  | :: |          || :||   |  |  | :| |: :||:||
Db        755 -HELEDEIIFITCHEGSLMALGFLIGYTCLLAAICFFFAFKSRKLPENFNEAKFITFSML 813

Qy        770 AYFITWVSFVPLLANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPE 827
              :|| |:||:|  |:       ||:: |:|    |:||     : |::: :|  || |
Db        814 IFFIVWISFIPAYASTYGKFVSAVEVIAILAASFGLLACIFFNKVYIILFKPSRNTIE 871


RESULT 4
B56715
calcium receptor (clone phPCaR-5.2) - human
C;Species: Homo sapiens (man)
C;Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 21-Jul-2000
C;Accession: B56715
R;Garrett, J.E.; Capuano, I.V.; Hammerland, L.G.; Hung, B.C.P.; Brown, E.M.; Hebert, S.C
J. Biol. Chem. 270, 12919-12925, 1995
A;Title: Molecular cloning and functional expression of human parathyroid calcium recepto
A;Reference number: A56715; MUID:95279439; PMID:7759551
A;Accession: B56715
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1088 <GAR>
A;Cross-references: UNIPARC:UPI000004A999; GB:U20760; NID:g683746; PIDN:AAA86504.1; PID:g
C;Keywords: glycoprotein; receptor; transmembrane protein

  Query Match             23.4%;  Score 1057.5;  DB 2;  Length 1088;
  Best Local Similarity   30.2%;  Pred. No. 3.8e-74;
  Matches  271;  Conservative  155;  Mismatches  365;  Indels  107;  Gaps   27;

Qy         12 WALL----HPGTGAPLCLSQQLRMKGDYVLGGLFPLG---EAEEAGLRSRTRPSSPVCTR 64
              | ||    |      |    |: : ||| :||||||:     |::  |:|  || |  | |
Db          8 WVLLALTWHTSAYGP---DQRAQKKGDIILGGLFPIHFGVAAKDQDLKS--RPESVECIR 62

Qy         65 FSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLA--KAGS 122
              ::  |  |  ||  |:||||:   ||| | ||| :||||:    |:: :| |:|  | |
Db         63 YNFRGFRWLQAMIFAIEEINSSPALLPNLTLGYRIFDTCNTVSKALEATLSFVAQNKIDS 122

Qy        123 RDIAAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETFPS 182
              ::  :|| ::: |  :||:|   | ::          | :|||||  :|  ||| :  | |
Db        123 LNLDEFCNCSEHIPSTIAVVGATGSGVSTAVANLLGLFYIPQVSYASSSRLLSNKNQFKS 182

Qy        183 FFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHEGL 242
              | ||:|:|  | || |::::  | ||||  : :||:||| |:  |   |  | |||    |
Db        183 FLRTIPNDEHQATAMADIIEYFRWNWVGTIAADDDYGRPGIEKFREEAEERDICIDFSEL 242

Qy        243 VPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVASEA 302
              :    :  |   ::| |:  :  |: :|:::|:|       |    :   :: ::|:||||
Db        243 I----SQYSDEEEIQHVVEVIQNSTAKVIVVFSSGPDLEPLIKEIVRRNITGKIWLASEA 298

Qy        303 WLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYV-KTHLALATDPAFCSALGERE--- 358
              | :| |:        :|  :||  :  |:  | ::: | |   :     |
Db        299 WASSSLIAMPQYFHVVGGTIGFALKAGQIPGFREFLKKVHPRKSVHNGFAKEFWEETFNC 358

Qy        359 ---------------QGLEEDVVGQRCPQCD------CITLQNVSA------GLNHHQ- 389
                            :| ||   | |             |   :|:|:          | :
Db        359 HLQEGAKGPLPVDTFLRGHEES--GDRFSNSSTAFRPLCTGDENISSVETPYIDYTHLRI 416

Qy        390 TFSVYAAVYSVAQALHNTLQC--------NASGCPAQDPVKPWQLLENMYNLTF-HVGGL 440
              :::|| ||||:| || :   |        | | |     |: ||:|::: :| | :  |
Db        417 SYNVYLAVYSIAHALQDIYTCLPGRGLFTNGS-CADIKKVEAWQVLKHLRHLNFTNNMGE 475

Qy        441 PLRFDSSGNVDMEYDLKLW---VWQGSVPRLHDVGRFN-GSLRTERL-----KIRWH--- 488
```

RESULT 5
I59362
calcium/polyvalent cation-sensing receptor precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I59362; A55594
R;Ruat, M.; Molliver, M.E.; Snowman, A.M.; Snyder, S.H.
Proc. Natl. Acad. Sci. U.S.A. 92, 3161-3165, 1995
A;Title: Calcium sensing receptor: molecular cloning in rat and localization to nerve te
A;Reference number: I59362; MUID:95241465; PMID:7724534
A;Accession: I59362
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1079 <RES>
A;Cross-references: UNIPROT:P48442; UNIPARC:UPI000012700E; EMBL:U20289; NID:g790578; PID
A;Experimental source: striatal
R;Riccardi, D.; Park, J.; Lee, W.S.; Gamba, G.; Brown, E.M.; Hebert, S.C.
Proc. Natl. Acad. Sci. U.S.A. 92, 131-135, 1995
A;Title: Cloning and functional expression of a rat kidney extracellular calcium/polyval
A;Reference number: A55594; MUID:95116508; PMID:7816802
A;Accession: A55594
A;Molecule type: mRNA
A;Residues: 1-133,'X',135-1079 <RIC>
A;Cross-references: UNIPARC:UPI000017C8D3; GB:U10354
A;Experimental source: kidney
C;Keywords: calcium; glycoprotein; phosphoprotein; transmembrane protein
F;1-20/Domain: signal sequence #status predicted <SIG>
F;187-212/Region: hydrophobic
F;613-635/Domain: transmembrane #status predicted <TM1>
F;650-670/Domain: transmembrane #status predicted <TM2>
F;683-700/Domain: transmembrane #status predicted <TM3>
F;725-744/Domain: transmembrane #status predicted <TM4>
F;770-790/Domain: transmembrane #status predicted <TM5>
F;806-828/Domain: transmembrane #status predicted <TM6>
F;841-860/Domain: transmembrane #status predicted <TM7>
F;90,261,287,386,468,488,594,893,1005/Binding site: carbohydrate (Asn) (covalent) #statu
F;794/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted
F;899,901/Binding site: phosphate (Ser) (covalent) (by protein kinase A) #status predict

Query Match 23.3%; Score 1054; DB 2; Length 1079;
Best Local Similarity 30.3%; Pred. No. 7.1e-74;
Matches 269; Conservative 163; Mismatches 363; Indels 92; Gaps 28;





RESULT 6
JH0561
metabotropic glutamate receptor 2 precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C;Accession: JH0561
R;Tanabe, Y.; Masu, M.; Ishii, T.; Shigemoto, R.; Nakanishi, S.
Neuron 8, 169-179, 1992
A;Title: A family of metabotropic glutamate receptors.
A;Reference number: JH0561; MUID:92110002; PMID:1309649
A;Accession: JH0561
A;Molecule type: mRNA
A;Residues: 1-872 <TAN>
A;Cross-references: UNIPROT:P31421; UNIPARC:UPI000012F07F

```
A;Experimental source: brain
C;Comment: This protein is coupled to a G protein and evokes a variety of functions by m
C;Superfamily: metabotropic glutamate receptor 4
C;Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane prot
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-872/Product: metabotropic glutamate receptor 2 #status predicted <MET>
F;568-590/Domain: transmembrane #status predicted <TRI>
F;605-625/Domain: transmembrane #status predicted <TII>
F;637-655/Domain: transmembrane #status predicted <III>
F;680-700/Domain: transmembrane #status predicted <TIV>
F;726-747/Domain: transmembrane #status predicted <TRV>
F;761-782/Domain: transmembrane #status predicted <TVI>
F;795-819/Domain: transmembrane #status predicted <VII>
F;203,286,338,402,547/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;601,675,827,837,843/Binding site: phosphate (Ser) (covalent) #status predicted
F;832/Binding site: phosphate (Thr) (covalent) #status predicted

  Query Match            14.8%;  Score 670.5;  DB 2;  Length 872;
  Best Local Similarity  27.8%;  Pred. No. 4.7e-44;
  Matches  256;  Conservative  133;  Mismatches  385;  Indels  147;  Gaps   37;

Qy          1 MLGPAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPL----GEAEEAGLRSRTR 56
              :||   | | ||  :  |       : | ::|| |||||||:    | ||| |  : |
Db          4 LLGFLAL-LLLWGAVAEGPA-----KKVLTLEGDLVLGGLFPVHQKGGPAEECGPVNEHR 57

Qy         57 PSSPVCTRFSSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMF 116
                        |:    ||  |:: ||    ||||:|||  : |:||:   |:: :| |
Db         58 -----------GIQRLEAMLFALDRINRDPHLLPGVRLGAHILDSCSKDTHALEQALDF 105

Qy        117 LAKA------GSRDI---AAYCNYTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSY 167
              :  :      ||| |    :|  ::     |  |||   |:::         | :||:||
Db        106 VRASLSRGADGSRHICPDGSYATHSDAPTAVTGVIGGSYSDVSIQVANLLRLFQIPQISY 165

Qy        168 GASMELLSARETFPSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFS 227
               ::   || :  :  | |||| |  |  | ||:|: | | :|: : |: :||  |:  |
Db        166 ASTSAKLSDKSRYDYFARTVPPDFFQAKAMAEILRFFNWTYVSTVASEGDYGETGIEAFE 225

Qy        228 ALAAARGICIAHEGLVPLPRADDSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYS 287
              | || ||:|    |    :  :  | |: :|    : | :| :||      | |   :
Db        226 LEARARNICVATSEKVGRAMSRAAFEGVVRALL---QKPSARVAVLFTRSEDARELL--A 280

Qy        288 ISSRLSPK-VWVASEAWLTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFP--QYVKTHLAL 344
              : ||:    ||||: |   : |:       | | :          :| :|    :     :|
Db        281 ATQRLNASFTWVASDGWGALESVVAGSERAAEGAIT------IELASYPISDFASYFQSL 334

Qy        345 -----ATDPAFCSALGEREQGLEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYS 399
                  : :| |       ||   |      |    :|: |           | |  |||:
Db        335 DPWNNSRNPWF------REFWEERFHCSFRQRDCAAHSLRAVPFEQESKIMF-VVNAVYA 387

Qy        400 VAQALHN---TLQCNASG-CPAQDPVKPWQLLEN-MYNLTFHVGGLP------LRFDSSG 448
              :| ||||    |  | :  | |  ||   :| :: : |: |      |       :|||  |
Db        388 MAHALHNMHRALCPNTTHLCDAMRPVNGRRLYKDFVLNVKFDAPFRPADTDDEVRFDRFG 447

Qy        449 NVDMEYDLKLWVWQGSVP-RLHDVGRFNGSLRTERLKIRWHT-SDNQKPVSRCSRQCQEG 506
              :    |::  ::  ||    |  || :   |  :    | | : |    | ||||  | :
Db        448 DGIGRYNIFTYLRAGSGRYRYQKVGYWAEGLTLDTSFIPWASPSAGPLPASRCSEPCLQN 507

Qy        507 QVRRVKGFHSCCYDCVDCEAGSYRQNPDDIACTFCGQDEWSPERSTRCFRRRSRFLAWGE 566
              :|: |:    ||: |: |:   ||   |:  |  ||    |     | ||      :: ||:
Db        508 EVKSVQPGEVCCWLCIPCQPYEYRL--DEFTCADCGLGYWPNASLTGCFELPQEYIRWGD 565

Qy        567 PAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQASGGPLACFGLVCLGLVCL---SVLL 623
                :  : :  |      |  ||:|| |  :|:|:|||  | |:  : || | |      :
Db        566 AWAVGPVTIACLGALATLFVLGVFVRHNATPVVKASGREL-CY--ILLGGVFLCYCMTFV 622

Qy        624 FPGQPSPARCLAQQPLSHLPLTGCLSTLFLQA---AEIF------VESELPLSWADRLSG 674
              |  :|| | |  ::       : | | |  :    | ||       :    :| | :::
Db        623 FIAKPSTAVCTLRRLGLGTAFSVCYSALLTKTNRIARIFGGAREGAQRPRFISPASQVAI 682

Qy        675 CLRGPWAWLVVLLAMLVEVALCTWYLVAFP----------PEVVTDWHMLPTEALVHCRT 724
              |       | ::   |: ||   | :|  |           ||||          : |
Db        683 C-------LALISGQLLIVA--AWLVVEAPGTGKETAPERREVVT----------LRCNH 723

Qy        725 RSWVSFGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVPLLAN 784
              |     | : | |  |  || |  |  |  | :| |: : | |     | |::|:|:
Db        724 RDASMLG-SLAYNVLLIALCTLYAFKTRKCPENFNEAKFIGFTMYTTCIIWLAFLPIF-- 780

Qy        785 VQVVLRPAVQMGALLLC---------VLGILAAFHLPRCYLLMRQPGLN-------TPEF 828
                  |     ::    :|         ||| | |   |: :::: ||  |       | |
Db        781 --YVTSSDYRVQTTTMCVSVSLSGSVVLGCLFA---PKLHIILFQPQKNVVSHRAPTSRF 835

Qy        829 FLGGGPGDAQGQNDGNTGNQG 849
                    | |  :   | | ||
Db        836 ------GSAAPRASANLG-QG 849


RESULT 7
A46742
metabotropic glutamate receptor, mGluR6 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A46742
R;Nakajima, Y.; Iwakabe, H.; Akazawa, C.; Nawa, H.; Shigemoto, R.; Mizuno, N.; Nakanishi,
J. Biol. Chem. 268, 11868-11873, 1993
A;Title: Molecular characterization of a novel retinal metabotropic glutamate receptor mG
A;Reference number: A46742; MUID:93280152; PMID:8389366
A;Accession: A46742
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-871 <NAK>
A;Cross-references: UNIPROT:P35349; UNIPARC:UPI000012F083; GB:D13963; NID:g391856; PIDN:
A;Experimental source: retina
A;Note: sequence extracted from NCBI backbone (NCBIN:133246, NCBIP:133250)
C;Superfamily: metabotropic glutamate receptor 4
C;Keywords: G protein-coupled receptor; transmembrane protein

  Query Match            14.0%;  Score 631.5;  DB 2;  Length 871;
  Best Local Similarity  25.6%;  Pred. No. 5.2e-41;
  Matches  225;  Conservative  146;  Mismatches  410;  Indels   99;  Gaps   31;

Qy          4 PAVLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSPVCT 63
              | :|    | |   |     | :  :|: |   ||||||:     ||           |
Db          5 PVLLLWLAWWLSQAGIA---CGAGSVRLAGGLTLGGLFPVHARGAAG---------RACG 52

Qy         64 RF-SSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMF----LA 118
                    |:    ||  |:: :|   :||||:|||  | ||||     |:: :| |    :
Db         53 ALKKEQGVHRLEAMLYALDRVNADPELLPGVRLGARLLDTCSRDTYALEQALSFVQALIR 112

Qy        119 KAGSRDIAAYCNYTQYQP-------RVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASM 171
              |  | |:          |      ||:||:|  :| ::::      | :||:|| ::
Db        113 GRGDGDEASVRCPGGVPPLRSAPPERVVAVVGASASSVSIMVANVLRLFAIPQISYASTA 172

Qy        172 ELLSARETFPSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALA- 230
              ||     :  | | || |  |  |  :::: |||:|: | |:  ||  |:  |  ::
Db        173 PELSDSTRYDFFSRVVPPDSYQAQAMVDIVRALGWNYVSTLASEGNYGESGVEAFVQISR 232

Qy        231 AARGICIAHEGLVPLPRADDSRLGKVQDVLHQVNQS-SVQVVLLFASVHAAHALFNYSIS 289
              | |:|||    : :||  : : |:    |: :: :: : : :::||:       :   :
Db        233 EAGGVCIAQS--IKIPR--EPKPGEPHKVIRRLMETPNARGIIIFANEDDIRRVLEATRQ 288

Qy        290 SRLSPK-VWVASEAWLTS-DLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTH------ 341
              : |:   :|| |::| :    ::  |   | :|  :   | : | :  | ||  |
Db        289 ANLTGHFLWVGSDSWGSKISPILNLEEEA-VGAIT-ILPKRASIDGFDQYFMTRSLENNR 346

Qy        342 ----LALATDPAF----CSALGEREQGLE----EDVVGQRCPQCDCITLQNVSAGLNHHQ 389
                  |   :  |      |: |: :         |: :||              ||     :
Db        347 RNIWFAEFWEENFNCKLTSSGGQSDDSTRKCTGEERIGQD-----------SAYEQEGK 394

Qy        390 TFSVYAAVYSVAQALHNTLQCNASG----CPAQDPVKPWQLLENMYNLTFH-VGGLPLRF 444
              |  |||::| |||:  |    |    ||| :|      ||  :  : |:   | |: |
Db        395 VQFVIDAVYAIAHALHSMHQALCPGHTGLCPAMEPTDGRTLLHYIRAVRFNGSAGTPVMF 454

Qy        445 DSSGNVDMEYDL-KLWVWQGSVPR--LHDVGRFNGSLRTERLKIRWHTSDNQKPVSRCSR 501
              : :|:    ||: :      ||       ||::  :|| :   :||    :: | |:||
Db        455 NENGDAPGRYDIFQYQATNGSASSGGYQAVGQWAEALRLDMEVLRWSGDPHEVPPSQCSL 514
```

```
Qy        502 QCQEGQVRR-VKGFHSCCYDCVDCEAGSYRQNPDDIACTFCGQDEWSPERSTRCFRRRSR 560
Db        515 PCGPGERKKMVKGV-PCCWHCEACD--GYRFQVDEFTCEACPGDMRPTPNHTGCRPTPVV 571

Qy        561 FLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQASGGPLACFGLVCLGLVCLS 620
Db        572 RLTWSSPWAALPLLLAVLGIMATTTIMATFMRHNDTPIVRASGRELSYVLLTGIFLIYAI 631

Qy        621 VLLFPGQPSPARCLAQQPLSHLPLTGCLSTLFLQA---AEIFVESELPLSWADRLSGCLR 677
Db        632 TFLMVAEPCAAICAARRLLLGLGTTLSYSALLTKTNRIYRIFEQGKRSVTPPPFIS---- 687

Qy        678 GPWAWLVVL--LAMLVEVALCTWYLVAFPPEVVTDWHMLPT------EALVHCRTRSWVS 729
Db        688 -PTSQLVITFGLTSLQVVGVIAW-LGAQPPHSVIDYEEQRTVDPEQARGVLKCDMSDLSL 745

Qy        730 FGLAHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVPLLANV---- 785
Db        746 IGCL-GYSLLLMVTCTVYAIKARGVPETFNEAKPIGFTMYTTCIIWLAFVPIFFGTAQSA 804

Qy        786 -QVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLN 824
Db        805 EKIYIQTTTLTVSLSLSASVSLGMLYVPKTYVILFHPEQN 844


RESULT 8
JH0563
metabotropic glutamate receptor 4 precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C;Accession: JH0563; I58149
R;Tanabe, Y.; Masu, M.; Ishii, T.; Shigemoto, R.; Nakanishi, S.
Neuron 8, 169-179, 1992
A;Title: A family of metabotropic glutamate receptors.
A;Reference number: JH0561; MUID:92110002; PMID:1309649
A;Accession: JH0563
A;Molecule type: mRNA
A;Residues: 1-912 <TAN>
A;Cross-references: UNIPROT:P31423; UNIPARC:UPI000012F080
A;Experimental source: brain
R;O'Hara, P.J.; Sheppard, P.O.; Thogersen, H.; Venezia, D.; Haldeman, B.A.; McGrane, V.;
Neuron 11, 41-52, 1993
A;Title: The ligand-binding domain in metabotropic glutamate receptors is related to bac
A;Reference number: I58149; MUID:93332699; PMID:8338667
A;Accession: I58149
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-123,'R',125-912 <RES>
A;Cross-references: UNIPARC:UPI0000167AAE; GB:M90518; NID:g205400; PIDN:AAA93190.1; PID:
C;Comment: This protein is coupled to a G protein and evokes a variety of functions by m
C;Genetics:
A;Gene: GLUR4
C;Superfamily: metabotropic glutamate receptor 4
C;Keywords: G protein-coupled receptor; glycoprotein; neurotransmitter receptor; phospho
F;1-32/Domain: signal sequence #status predicted <SIG>
F;33-912/Product: metabotropic glutamate receptor 4 #status predicted <MET>
F;588-610/Domain: transmembrane #status predicted <TRI>
F;625-645/Domain: transmembrane #status predicted <TII>
F;657-675/Domain: transmembrane #status predicted <III>
F;700-720/Domain: transmembrane #status predicted <TIV>
F;751-772/Domain: transmembrane #status predicted <TRV>
F;786-807/Domain: transmembrane #status predicted <TVI>
F;822-847/Domain: transmembrane #status predicted <VII>
F;98,301,454,484,569/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;621,689,695,859,870/Binding site: phosphate (Ser) (covalent) #status predicted

  Query Match           13.7%;  Score 619.5;  DB 2;  Length 912;
  Best Local Similarity  24.6%;  Pred. No. 4.8e-40;
  Matches  222;  Conservative  145;  Mismatches  397;  Indels  137;  Gaps   31;

Qy          4 PAVLGLSLWALLHPGT-----GAPLCLSQQLRMKGDYVLGGLFPL-GEAEEAGLRSRTRP 57
Db         15 PLCLLLSLYAPWVPSSLGKPKGHPHMNS--IRIDGDITLGGLFPVHGRGSE--------- 63

Qy         58 SSPVCTRF-SSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMF 116
Db         64 -GKACGELKKEKGIHRLEAMLFALDRINNDPDLLPNITLGARILDTCSRDTHALEQSLTF 122

Qy        117 LAKAGSRD-IAAYCN-----YTQYQPRVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGAS 170
Db        123 VQALIEKDGTEVRCGSGGPPIITKPERVVGVIGASGSSVSIMVANILRLFKIPQISYAST 182

Qy        171 MELLSARETFPSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALA 230
Db        183 APDLSDNSRYDFFSRVVPSDTYQAQAMVDIVRALKWNYVSTLASEGSYGESGVEAFIQKS 242

Qy        231 AAR-GICIAHEGLVPLPRADDSRLGKVQDVLHQ-VNQSSVQVVLLFAS-------VHAAH 281
Db        243 RENGGVCIAQS--VKIPR--EPKTGEFDKIIKRLLETSNARGIIIFANEDDIRRVLEAAR 298

Qy        282 ALFNYSISSRLSPKVWVASEAW-LTSDLVMGLPGMAQMGTVLGFLQRGAQLHEFPQYVKT 340
Db        299 R------ANQTGHFFWMGSDSWGSKSAPVLRLEEVAE-GAVT-ILPKRMSVRGFDRYFSS 350

Qy        341 H-----------LALATDPAFCSALGEREQGLEEDVVGQRCPQCDCITLQNVSAGLNHHQ 389
Db        351 RTLDNNRRNIWFAEFWEDNFHCKL---SRHALKKGSHIKKCTNRERIGQD--SAYEQEGK 405

Qy        390 TFSVYAAVYSVAQALHNTLQCNASG----CPAQDPVKPWQLLENMYNLTFH-VGGLPLRF 444
Db        406 VQFVIDAVYAMGHALHAMHRDLCPGRVGLCPRMDPVDGTQLLKYIRNVNFSGIAGNPVTF 465

Qy        445 DSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLRTERLKIRWHTSDNQKPVSRCSRQCQ 504
Db        466 NENGDAPGRYDIYQYQLRNGSAEYKVIGSWTDHLHLRIERMQWPGSGQQLPRSICSLPCQ 525

Qy        505 EGQVRR-VKGFHSCCYDCVDCEAGSYRQNPDDIACTFCGQDEWSPERSTRCFRRRSRFLA 563
Db        526 PGERKKTVKGM-ACCWHCEPCT--GYQYQVDRYTCKTCPYDMRPTENRTSCQPIPIVKLE 582

Qy        564 WGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQASGGPLACFGLVCLGLVCLSVLL 623
Db        583 WDSPWAVLPLFLAVVGIAATLFVVVTFVRYNDTPIVKASGRELSYVLLAGIFLCYATTFL 642

Qy        624 FPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIF---------VESELPLSWADRLSG 674
Db        643 MIAEPDLGTCSLRRIFLGLGMSISYAALLTKTNRIYRIFEQGKRSVSAPRFISPASQLA- 701

Qy        675 CLRGPWAWLVVLLAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALVHCRTRSWVSFGLAH 734
Db        702 --------ITFILISLQLLGICVWFVVD-PSHSVVDFQDQRTLDPRFAR-------GVLK 745

Qy        735 ATNATLAFLCFLGTFLV-----------RSQPGCYNRARGLTFAMLAYFITWVSFVPLL 782
Db        746 CDISDLSLICLLGYSMLLMVTCTVYAIKTRGVPETFNEAKPIGFTMYTTCIVWLAFIPIF 805

Qy        783 ---------------ANVQVVLRPAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTP 826
Db        806 FGTSQSADKLYIQTTTLTVSVSLSASVSLGML----------YMPKVYIILFHPEQNVP 854

Qy        827 E 827
Db        855 K 855


RESULT 9
S71376
glutamate receptor homolog - cherry salmon
C;Species: Oncorhynchus masou (cherry salmon)
C;Date: 11-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 09-Jul-2004
C;Accession: S71376
R;Kubokawa, K.; Miyashita, T.; Nagasawa, H.; Kubo, Y.
FEBS Lett. 392, 71-76, 1996
A;Title: Cloning and characterization of a bifunctional metabotropic receptor activated
A;Reference number: S71376; MUID:96354880; PMID:8769318
A;Accession: S71376
```

```
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1218 <KUB>
A;Cross-references: UNIPROT:Q90ZF3; UNIPARC:UPI000017BF56
C;Keywords: glycoprotein; phosphoprotein
F;603-625/Domain: transmembrane #status predicted <TM1>
F;640-660/Domain: transmembrane #status predicted <TM2>
F;672-690/Domain: transmembrane #status predicted <TM3>
F;717-737/Domain: transmembrane #status predicted <TM4>
F;761-782/Domain: transmembrane #status predicted <TM5>
F;796-817/Domain: transmembrane #status predicted <TM6>
F;826-850/Domain: transmembrane #status predicted <TM7>
F;104,233,403,525,757/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;636,699,961/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status pre
F;705/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
F;892/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status predic

  Query Match           13.5%;  Score 612.5;  DB 2;  Length 1218;
  Best Local Similarity  25.7%;  Pred. No. 2.4e-39;
  Matches  226;  Conservative  140;  Mismatches  370;  Indels  143;  Gaps   34;

Qy         30 RMKGDYVLGGLFPL--------------GEAEEAGLRSRTRPSSPVCTRFSSNGLLWALA 75
              || || ::| || :                || |                  |:    |
Db         46 RMDGDIIIGALFSVHHQPSTEQVAERKCGEVRE-----------------QYGIQRVEA 87

Qy         76 MKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLA-----------KAGSRD 124
              |   :: ||:  :||| : || :: |:|     ||:: |: |:            | |:
Db         88 MFHTLDRINSNPNLLPNITLGCEIRDSCWHSSVALEQSIEFIRDSLISIRDDDNKDGTSR 147

Qy        125 IAAYCNYTQYQP-----RVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARET 179
                          ||       :  |||| || :|:         | :||::| |:   || :
Db        148 QWCIEGMPPSQPPATKRPIAGVIGPGSSSVAIQVQNLLQLFNIPQIAYSATSIDLSDKTL 207

Qy        180 FPSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAH 239
              |  | | |||| :|  |  :::: : | :|:|: ::  ||  |:  |  ||:  |:||||
Db        208 FKYFLRVVPSDTLQARAILDIVKRYNWTYVSAVHTEGNYGESGMEAFKELASQEGLCIAH 267

Qy        240 EGLVPLPRAD---DSRLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKV 296
                 :     :   | | |::: |         :||: |        |    :: |   ::
Db        268 SDKIYSNAGEKHFDRLLRKLRERL-----PKARVVVCFCEGMTVRGLL---MAMR---RL 316

Qy        297 WVASEAWLTSDLVMGLPGMAQMGTVL-GFLQRG-----AQLH-----EFPQY-VKTHLAL 344
              || |      |::|  | |    |: |: |        :||       |  | :|  |
Db        317 GVAGEF-----LLIGSDGWADRDEVVEGYEQEAVGGITVKLHSEEVTSFDDYFLKLRLNT 371

Qy        345 AT-DPAFCSALGEREQ------GLEEDVVGQRCPQCDCIT---LQNVSAGLNHHQTFSVY 394
              | :| |       | |       ||      : |   : :    :|:   |        |
Db        372 NTRNPWFPEFWQHRFQCRIPGHPLENMNYRKNCSGYESLEDNYVQDSKMGF-------VI 424

Qy        395 AAVYSVAQALHNTLQCNASG----CPAQDPVKPWQLLENMYNLTF-HVGGLPLRFDSSGN 449
              |:|::|| ||:      |    | | ||:   |||| :   :|  | |  : || :|:
Db        425 NAIYAMAQGLHDMHSHLCPGHVGLCKAMDPIDGSQLLEFLMRTSFTGVSGEDVWFDENGD 484

Qy        450 VDMEYDLK--LWVWQGSVPRLHDVGRFNGSLRTERLKIRWHTSDNQKPVSRCSRQCQEGQ 507
                 |::    :|  |:   ::       | |  :   :: : ||    :| ||  | :|:
Db        485 TPGRYEIMNLQYVEPGAFDYINVGSWHEGQLSIDDYMMQINRSD--MVLSVCSEPCSKGE 542

Qy        508 VRRV-KGFHSCCYDCVDCEAGSYRQNPDDIACTFCGQDEWSPERSTRCFRRRSRFLAWGE 566
              :: : ||  |||: |  |:       |  |:  || |    |        |      |:| ||
Db        543 IKVIRKGEVSCCWICTACKDNEIVQ--DEFTCTACDLGWWPDPELEGCEPITLRYLEWGN 600

Qy        567 PAVLLLLLLLSLALGLVLAALGLFVHHRDSPLVQASGGPLACFGL---VCLGLVCLSVLL 623
              |  :: ::    | : :       :|| :||:|:|::|    | |: :   : || :|   |:
Db        601 PESIVQVVFACLGILVTSFVTFIFVLYRDTPVVKSSSREL-CYIILAGIFLGYICPFTLI 659

Qy        624 FPGQPSPARCLAQQPLSHLPLTGCLSTLFLQAAEIFVESELPLSWADRLSG-----CLRG 678
                 ||: | |  |: |  |  | | | |  :   |           | |:|     | |
Db        660 --AQPTVASCYLQRLLVGLSATMCYSALVTKTNRI----------ARILAGSKKKICTRK 707

Qy        679 P-----WAWLVVL-LAMLVEVALCTWYLVAFPPEVVTDWHMLPTEALVHCRTRSWVSFGL 732
              |     || ||:  | : |:: |    ::  ||  |  :  :  |  : | | | |
Db        708 PRFMSAWAQLVIAGLLVSVQLTLEVTLIILEPPMPVKSYPSI-REVFLICNT-STVGMVA 765

Qy        733 AHATNATLAFLCFLGTFLVRSQPGCYNRARGLTFAMLAYFITWVSFVPLL--ANVQVVLR 790
                 |  |   |   |  |: |  :| |: : | |    | |::|||:   :| :::
Db        766 PLGYNGLLIMSCTYYAFKTRNVPANFNEAKYIAFTMYTTCIIWLAFVPIYFGSNYKII-- 823

Qy        791 PAVQMGALLLCVLGILAAFHLPRCYLLMRQPGLNTPEFF 829
                    :: | |   |      |: |::: :|  |    |
Db        824 --TTSFSVSLSVTVALGCMFSPKIYIILAKPERNVRSAF 860


RESULT 10
A49874
metabotropic glutamate receptor 7 - rat
N;Alternate names: metabotropic glutamate receptor mGluR7
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: A49874; I57954
R;Okamoto, N.; Hori, S.; Akazawa, C.; Hayashi, Y.; Shigemoto, R.; Mizuno, N.; Nakanishi,
J. Biol. Chem. 269, 1231-1236, 1994
A;Title: Molecular characterization of a new metabotropic glutamate receptor mGluR7 coupl
A;Reference number: A49874; MUID:94117433; PMID:8288585
A;Accession: A49874
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-915 <RES>
A;Cross-references: UNIPROT:P35400; UNIPARC:UPI000012F084; GB:D16817; NID:g458728; PIDN:
R;Saugstad, J.A.; Kinzie, J.M.; Mulvihill, E.R.; Segerson, T.P.; Westbrook, G.L.
Mol. Pharmacol. 45, 367-372, 1994
A;Title: Cloning and expression of a new member of the L-2-amino-4-phosphonobutyric acid-
A;Reference number: I57954; MUID:94195260; PMID:8145723
A;Accession: I57954
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-915 <RE2>
A;Cross-references: UNIPARC:UPI000012F084; EMBL:U06832; NID:g459657; PIDN:AAA20655.1; PI
C;Genetics:
A;Gene: MGLUR7
C;Superfamily: metabotropic glutamate receptor 4
C;Keywords: neurotransmitter receptor

  Query Match           13.4%;  Score 604.5;  DB 2;  Length 915;
  Best Local Similarity  23.9%;  Pred. No. 7.1e-39;
  Matches  212;  Conservative  169;  Mismatches  386;  Indels  121;  Gaps   32;

Qy          6 VLGLSLWALLHPGTGAPLCLSQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSPVCTRF 65
              || : |  |      |  :     :|::||  ||||||:     :|:        |
Db         20 VLEVLLCVLAAAARGQEMYAPHSIRIEGDVTLGGLFPVHAKGPSGV--------PCGDIK 71

Qy         66 SSNGLLWALAMKMAVEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKAGSRDI 125
                ||:    ||  |:::||:  :||| : ||  : ||||    |:: || |:      :|
Db         72 RENGIHRLEAMLYALDQINSDPNLLPNVTLGARILDTCSRDTYALEQSLTFVQALIQKDT 131

Qy        126 A-AYCNYTQYQP-------RVLAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSAR 177
              :   |  |  :|       :|: |||    | ::::       | :||:|| ::   ||
Db        132 SDVRC--TNGEPPVFVKPEKVVGVIGASGSSVSIMVANILRLFQIPQISYASTAPELSDD 189

Qy        178 ETFPSFFRTVPSDRVQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALA-AARGIC 236
              :  | | || |  |  |  ::::  |||:|: | |:  || :|:  |: ::  | |:|
Db        190 RRYDFFSRVVPPDSFQAQAMVDIVKALGWNYVSTLASEGSYGEKGVESFTQISKEAGGLC 249

Qy        237 IAHEGLVPLPRADDSRLGKVQDVLHQ-VNQSSVQVVLLFAS-VHAAHALFNYSISSRLSP 294
              ||    :|  | |  |       :: | ::  :  : |::||:       |      : ::
Db        250 IAQSVRIPQERKD--RTIDFDRIIKQLLDTPNSRAVVIFANDEDIKQILAAAKRADQVGH 307

Qy        295 KVWVASEAWLT--------SDLVMG-LPGMAQMGTVLGF--------LQRGAQLHEFPQY 337
              :|| |::| :         |:  | :    : || ||          |:   :   | :|
Db        308 FLWVGSDSWGSKINPLHQHEDIAEGAITIQPKRATVEGFDAYFTSRTLENNRRNVWFAEY 367

Qy        338 VKTHLALATDPAFCSALGEREQGLEEDVVGQRCPQCDCITLQNVSAGLNHHQTFS---VY 394
              : :           :  | :::  :    ||          : :    |: |      |
Db        368 WEENFNCK-----LTISGSKKEDTDRKCTGQ----------ERIGKDSNYEQEGKVQFVI 412

Qy        395 AAVYSVAQALHN---TLQCNASG-CPAQDPVKPWQLLENMYNLTFH-VGGLPLRFDSSGN 449
              |||::| |||:    |  :  | ||  :     :||: : :: |:   | |: |: :|:
```

```
Db        413 DAVYAMAHALHHMNKDLCADYRGVCPEMEQAGGKKLLKYIRHVNFNGSAGTPVMFNKNGD 472

Qy        450 VDMEYDLKLWVWQGSV-PRLHDVGRFNGSLRTERLKIRWHTSDNQKPVSRCSRQCQEGQV 508

Db        473 APGRYDIFQYQTTNTTNPGYRLIGQWTDELQLNIEDMQWGKGVREIPSSVCTLPCKPGQR 532

Qy        509 RRVKGFHSCCYDCVDCEAGSYRQNPDDIACTFCGQDEWSPERSTRCFRRRSRFLAWGEPA 568

Db        533 KKTQKGTPCCWTCEPCD--GYQYQFDEMTCQHCPYDQRPNENRTGCQNIPIIKLEWHSPW 590

Qy        569 VLLLLLLLSLALGLVLAALGLFVHHRDSPLVQASGGPLACFGLVCLGLVCLSVLLFPGQP 628

Db        591 AVIPVFLAMLGIIATIFVMATFIRYNDTPIVRASGRELSYVLLTGIFLCYIITFLMIAKP 650

Qy        629 SPARCLAQQPLSHLPLTGCLSTLFL-----QAAEIFVESELPLSWADRLSGCLRGPWAWL 683

Db        651 DVAVCSFRRVF--LGLGMCISYAALLTKTNRIYRIFEQGKKSVT-APRLI----SPTSQL 703

Qy        684 VVLLAMLVEVALC---TWYLVAFPPEVVTDWH----MLPTEALVHCRTRSWVSFGLAHAT 736

Db        704 AI-TSSLISVQLLGVFIWFGVD-PPNIIIDYDEHKTMNPEQAR-----------GVLKCD 750

Qy        737 NATLAFLCFLGTFLV------------RSQPGCYNRARGLTFAMLAYFITWVSFVPL-LA 783

Db        751 ITDLQIICSLGYSILLMVTCTVYAIKTRGVPENFNEAKPIGFTMYTTCIVWLAFIPIFFG 810

Qy        784 NVQVVLRPAVQMGALLLCV-------LGILAAFHLPRCYLLMRQPGLN 824

Db        811 TAQSAEKLYIQTTTLTISMNLSASVALGML---YMPKVYIIIFHPELN 855


RESULT 11
JH0562
metabotropic glutamate receptor 3 precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C;Accession: JH0562
R;Tanabe, Y.; Masu, M.; Ishii, T.; Shigemoto, R.; Nakanishi, S.
Neuron 8, 169-179, 1992
A;Title: A family of metabotropic glutamate receptors.
A;Reference number: JH0561; MUID:92110002; PMID:1309649
A;Accession: JH0562
A;Molecule type: mRNA
A;Residues: 1-879 <TAN>
A;Cross-references: UNIPROT:P31422; UNIPARC:UPI000004E90A
A;Experimental source: brain
C;Comment: This protein is coupled to a G protein and evokes a variety of functions by m
C;Superfamily: metabotropic glutamate receptor 4
C;Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane prot
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-879/Product: metabotropic glutamate receptor 3 #status predicted <MET>
F;577-599/Domain: transmembrane #status predicted <TRI>
F;614-634/Domain: transmembrane #status predicted <TII>
F;646-664/Domain: transmembrane #status predicted <III>
F;689-709/Domain: transmembrane #status predicted <TIV>
F;735-756/Domain: transmembrane #status predicted <TRV>
F;770-791/Domain: transmembrane #status predicted <TVI>
F;804-828/Domain: transmembrane #status predicted <VII>
F;209,292,414,439/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;610,845/Binding site: phosphate (Ser) (covalent) #status predicted

  Query Match             13.2%;  Score 597.5;  DB 2;  Length 879;
  Best Local Similarity   25.2%;  Pred. No. 2.4e-38;
  Matches  218;  Conservative  142;  Mismatches  380;  Indels  125;  Gaps   32;

Qy         25 LSQQLRMKGDYVLGGLFPLGE----AEEAGLRSRTRPSSPVCTRFSSNGLLWALAMKMAV 80

Db         29 MRREIKIEGDLVLGGLFPINEKGTGTEECGRINEDR------------GIQRLEAMLFAI 76

Qy         81 EEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFL-AKAGSRDIAAY-CNYTQYQ--- 135

Db         77 DEINKDNYLLPGVKLGVHILDTCSRDTYALEQSLEFVRASLTKVDEAEYMCPDGSYAIQE 136

Qy        136 --PRVLA-VIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETFPSFFRTVPSDRV 192

Db        137 NIPLLIAGVIGGSYSSVSIQVANLLRLFQIPQISYASTSAKLSDKSRYDYFARTVPPDFY 196

Qy        193 QLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHEGLVPLPRADDSR 252

Db        197 QAKAMAEILRFFNWTYVSTVASEGDYGETGIEAFEQEARLRNICIATAEKVGRSNIRKSY 256

Qy        253 LGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPK-VWVASEAW-LTSDLVM 310

Db        257 DSVIRELLQKPN---ARVVVLFMRSDDSRELI--AAANRVNASFTWVASDGWGAQESIVK 311

Qy        311 GLPGMAQMGTVLGFLQRGAQLHEFPQYVKT------HLALATDPAF---------CSALG 355

Db        312 GSEHVAYGAITLELASH--PVRQFDRYFQSLNPYNNH----RNPWFRDFWEQKFQCSLQN 365

Qy        356 EREQGLEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALH---NTLQCNA 412

Db        366 KRNH-------RQVCDKHLAIDSSNYE---QESKIMFVVNAVYAMAHALHKMQRTLCPNT 415

Qy        413 SG-CPAQDPVKPWQLL-ENMYNLTFHVGGLP-------LRFDSSGNVDMEYDLKLWVWQG 463

Db        416 TKLCDAMKILDGKKLYKEYLLKINFTAPFNPNKGADSIVKFDTFGDGMGRYNVFNLQQTG 475

Qy        464 SVPRLHDVGRFNGSLRTERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCVD 523

Db        476 GKYSYLKVGHWAETLSLDVDSIHW--SRNSVPTSQCSDPCAPNEMKNMQPGDVCCWICIP 533

Qy        524 CEAGSYRQNPDDIACTFCGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLV 583

Db        534 CE--PYEYLVDEFTCMDCGPGQWPTADLSGCYNLPEDYIKWEDAWAIGPVTIACLGFLCT 591

Qy        584 LAALGLFVHHRDSPLVQASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLP 643

Db        592 CIVITVFIKHNNTPLVKASGRELCYILLFGVSLSYCMTFFFIAKPSPVICALRRLGLGTS 651

Qy        644 LTGCLSTLFLQA---AEIFVESELPLSWADRLSGCLRG--------PWAWLVVLLAM-LV 691

Db        652 FAICYSALLTKTNCIARIF------------DGVKNGAQRPKFISPSSQVFICLGLILV 698

Qy        692 EVALCTWYLVAFPPEVVTDWHMLPTE---ALVHCRTRSWVSFGLAHATNATLAFLCFLGT 748

Db        699 QIVMVSVWLILETPG--TRRYTLPEKRETVILKCNVKD-SSMLISLTYDVVLVILCTVYA 755

Qy        749 FLVRSQPGCYNRARGLTFAMLAYFITWVSFVPLLANVQVVLRPAVQMGALLLC------- 801

Db        756 FKTRKCPENFNEAKFIGFTMYTTCIIWLAFLPIF----YVTSSDYRVQTTTMCISVSLSG 811

Qy        802 --VLGILAAFHLPRCYLLMRQPGLN 824

Db        812 FVVLGCLFA---PKVHIVLFQPQKN 833


RESULT 12
I49142
metabotropic glutamate receptor 8 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I49142
R;Duvoisin, R.M.; Zhang, C.; Ramonell, K.
J. Neurosci. 15, 3075-3083, 1995
A;Title: A novel metabotropic glutamate receptor expressed in the retina and olfactory b
A;Reference number: I49142; MUID:95239344; PMID:7722646
A;Accession: I49142
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-908 <RES>
A;Cross-references: UNIPROT:P47743; UNIPARC:UPI00000274BE; EMBL:U17252; NID:g854728; PID
C;Genetics:
A;Gene: mGluR8
C;Superfamily: metabotropic glutamate receptor 4
C;Keywords: neurotransmitter receptor
```

Query Match 13.1%; Score 593.5; DB 2; Length 908;
Best Local Similarity 23.4%; Pred. No. 5.1e-38;
Matches 206; Conservative 157; Mismatches 368; Indels 149; Gaps 29;

```
Qy         26 SQQLRMKGDYVLGGLFPLGEAEEAGLRSRTRPSSPVCTRFSSNGLLWALAMKMAVEEINN 85
              :  :|: || :||||||:    | |:         |         |:    ||  |::: |
Db         37 AHSIRLDGDIILGGLFPVHAKGERGV--------PCGDLKKEKGIHRLEAMLYAIDQTNK 88

Qy         86 KSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFLAKAGSRDIAAYCNYTQYQP-------RV 138
                 |||  : ||  : ||||     |:: || |:     :| |:          |        ::
Db         89 DPDLLSNITLGVRILDTCSRDTYALEQSLTFVQALIEKD-ASDVKCANGDPPIFTKPDKI 147

Qy        139 LAVIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETFPSFFRTVPSDRVQLTAAA 198
                 |||  :| :::::       | :||:|| ::   ||    :  | | || |  |  |
Db        148 SGVIGAAASSVSIMVANILRLFKIPQISYASTAPELSDNTRYDFFSRVVPPDSYQAQAMV 207

Qy        199 ELLQEFGWNWVAALGSDDEYGRQGLSIFSALA-AARGICIAHEGLVPLPRADDSRLGKVQ 257
              :::   |||:|: | |:  ||  |:  |: ::    |:|||    :|     : | |: :
Db        208 DIVTALGWNYVSTLASEGNYGESGVEAFTQISREIGGVCIAQSQKIP----REPRPGEFE 263

Qy        258 DVLHQVNQS-SVQVVLLFAS-------VHAAHALFNYSISSRLSPKVWVASEAWLTSDLV 309
              :: :: :: : : |::||:       : ||  |      ::    :|: |::| :
Db        264 KIIKRLLETPNARAVIMFANEDDIRGILEAAKKL------NQSGHFLWIGSDSWGSKIAP 317

Qy        310 MGLPGMAQMGTVLGFLQRGAQLHEFPQYVKTH-LALATDPAFCSALGEREQGLEEDVVGQ 368
              :      | |   | : | :  | :| ::  ||      : :   |   | :    |:
Db        318 VYQQEEIAEGAVT-ILPKRASIDGFDRYFRSRTLANNRRNVWFAEFSEGNFGCKSGSHGK 376

Qy        369 RCPQC-DCITLQNVSAGLNHHQTFS---VYAAVYSVAQALHNTLQCNASG----CPAQDP 420
              |       |  |: ::   :: |       |  ||||:| ||||  :    |    ||
Db        377 RNSHIKKCTGLERIARDSSYEQEGKVQFVIDAVYSMAYALHNMHKELCPGYIGLCPRMVT 436

Qy        421 VKPWQLLENMYNLTFH-VGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHDVGRFNGSLR 479
              :   :|| :  : |:   | |: |: :|:    ||:  :          :| :   |
Db        437 IDGKELLGYIRAVNFNGSAGTPVTFNENGDAPGRYDIFQYQINNKSTEYKIIGHWTNQLH 496

Qy        480 TERLKIRWHTSDNQKPVSRCSRQCQEGQVRR-VKGFHSCCYDCVDCEAGSYRQNPDDIAC 538
              :  ::|   ::  | | ||  |: |: :: |||   ||: |  ||   |    |:::|
Db        497 LKVEDMQWANREHTHPASVCSLPCKPGERKKTVKGV-PCCWHCGRCE--GYNYQVDELSC 553

Qy        539 TFCGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGLVLAALGLFVHHRDSPL 598
              |  |:      | | |     | |  |  :: :|:  | :       :  || : |:|:
Db        554 ELCPLDQRPNINRTGCQRIPIIKLEWHSPWAVVPVLIAILGIIATTFVIVTFVRYNDTPI 613

Qy        599 VQASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHLPLTG---CLSTLFLQA 655
              |:|||                                : ||:: |||   | | ||
Db        614 VRASG--------------------------------RELSYVLLTGIFLCYSITFLMI 640

Qy        656 A----------EIFVESELPLSWADRLSGCLR-----------------GPWAWLVVLL 687
              |          ||:   :  |:|  |:   |                 | : ||:
Db        641 AAPDTIICSFRRIFLGLGMCFSYAALLTKTNRIHRIFEQGKKSVTAPKFISPASQLVITF 700

Qy        688 AMLVEVALC---TWYLVAFPPEVVTDWHMLPTEALVHCRTRSWVSFGLAHATNATLAFLC 744
              : |: | |      |::|  ||  : |:     |     : |         |:      : |: :|
Db        701 S-LISVQLLGVFVWFVVD-PPHTIIDYGEQRTLDPENAR-------GVLKCDISDLSLIC 751

Qy        745 FLGTFLV-----------RSQPGCYNRARGLTFAMLAYFITWVSFVPL-LANVQVVLRP 791
              ||  ::           | |  :| |: : | |     | |::|:|:     |   :
Db        752 SLGYSILLMVTCTVYAIKTRGVPETFNEAKPIGFTMYTTCIIWLAFIPIFFGTAQSAEKM 811

Qy        792 AVQMGALLLCV-------LGILAAFHLPRCYLLMRQPGLN 824
              :|   | : :       ||:|   ::|: |:::   |  |
Db        812 YIQTTTLTVSMSLSASVSLGML---YMPKVYIIIFHPEQN 848
```

RESULT 13
JC7160
metabotropic glutamate receptor subtype 3 precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 04-Mar-2000 #sequence_revision 04-Mar-2000 #text_change 11-May-2000
C;Accession: JC7160
R;Minoshima, T.; Nakanishi, S.
J. Biochem. 126, 889-896, 1999

A;Title: Structural organization of the mouse metabotropic glutamate receptor subtype 3 g
A;Reference number: JC7160; MUID:20012997; PMID:10544282
A;Accession: JC7160
A;Molecule type: DNA
A;Residues: 1-879 <MIN>
A;Cross-references: UNIPARC:UPI000004E90B; GB:AF170696
C;Genetics:
A;Gene: mGluR3
C;Superfamily: metabotropic glutamate receptor 4
C;Keywords: differentiation; G protein-coupled receptor; receptor; transmembrane protein
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-879/Product: metabotropic glutamate receptor subtype 3 #status predicted <MAT>
F;577-599/Domain: transmembrane #status predicted <TM1>
F;614-634/Domain: transmembrane #status predicted <TM2>
F;646-664/Domain: transmembrane #status predicted <TM3>
F;689-709/Domain: transmembrane #status predicted <TM4>
F;735-756/Domain: transmembrane #status predicted <TM5>
F;770-791/Domain: transmembrane #status predicted <TM6>
F;804-828/Domain: transmembrane #status predicted <TM7>

Query Match 13.1%; Score 592.5; DB 2; Length 879;
Best Local Similarity 25.1%; Pred. No. 5.8e-38;
Matches 217; Conservative 144; Mismatches 378; Indels 127; Gaps 33;

```
Qy         25 LSQQLRMKGDYVLGGLFPLGE----AEEA-GLRSRTRPSSPVCTRFSSNGLLWALAMKMA 79
              : :::::|| ||||||: |     ||  |:                    |:    ||  |
Db         29 MRREIKIEGDLVLGGLFPINEKGTGTEECRGIN------------EDRGIQRLEAMLFA 75

Qy         80 VEEINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLMFL-AKAGSRDIAAY-CNYTQYQ-- 135
              ::|||  : ||||::||  : ||||     |:: || |: |     | | | |     |
Db         76 IDEINKDNYLLPGVKLGVHILDTCSRDTYALEQSLEFVRASLTKVDEAEYMCPDGSYAIQ 135

Qy        136 ---PRVLA-VIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETFPSFFRTVPSDR 191
                 | ::| |||   | :::         | :||:|| ::   || :  :  | |||| |
Db        136 ENIPLLIAGVIGGSYSSVSIQVANLLRLFQIPQISYASTSAKLSDKSRYDYFARTVPPDF 195

Qy        192 VQLTAAAELLQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHEGLVPLPRADDS 251
              |  | ||:|: | | :|: : |: :||   |:  |   |  | ||||     |       |
Db        196 YQAKAMAEILRYFNWTYVSTVASEGDYGETGIEAFEQEARLRNICIATAEKVGRSNIRKS 255

Qy        252 RLGKVQDVLHQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPK-VWVASEAW-LTSDLV 309
                 ::::| : |    :||:||      :  |   : :||::    ||||: |     :|
Db        256 YDSVIRELLQKPN---ARVVVLFMRSDDSRELI--AAASRVNASFTWVASDGWGAQESIV 310

Qy        310 MGLPGMAQMGTVLGFLQRGAQLHEFPQYVKT------HLALATDPAF---------CSAL 354
              |   :|      |          : :| :| ::      |     :| |          ||
Db        311 KGSEHVAYGAITLELASH--PVRQFDRYFQSLNPYNNH----RNPWFRDFWEQKFQCSLQ 364

Qy        355 GEREQGLEEDVVGQRCPQCDCITLQNVSAGLNHHQTFSVYAAVYSVAQALH---NTLQCN 411
              :|          | |  :   |     |          :   |  |||::| |||    ||  |
Db        365 NKRNH-------RQICDKHLAIDSSNYE---QESKIMFVVNAVYAMAHALHKMQRTLCPN 414

Qy        412 ASG-CPAQDPVKPWQLLEN-MYNLTFHVGGLP-------LRFDSSGNVDMEYDLKLWVWQ 462
              :  | |   :  :| :: :  : |      |        ::||: |:    |::  :
Db        415 TTKLCDAMKILDGKKLYKDYLLKINFTAPFNPNKGADSIVKFDTYGDGMGRYNVFNFQHI 474

Qy        463 GSVPRLHDVGRFNGSLRTERLKIRWHTSDNQKPVSRCSRQCQEGQVRRVKGFHSCCYDCV 522
              |      || :  :|  :   | | | | | |:||  |   ::: ::     ||: |:
Db        475 GGKYSYLKVGHWAETLYLDVDSIHW--SRNSVPTSQCSDPCAPNEMKNMQPGDVCCWICI 532

Qy        523 DCEAGSYRQNPDDIACTFCGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGL 582
              ||   |    |:  |  ||  :|     : |:     :: | :   :  : :  |
Db        533 PCE--PYEYLVDEFTCMDCGPGQWPTADLSGCYNLPEDYIRWEDAWAIGPVTIACLGFMC 590

Qy        583 VLAALGLFVHHRDSPLVQASGGPLACFGLVCLGLVCLSVLLFPGQPSPARCLAQQPLSHL 642
                : :|: | ::|||:|||  |     |  : |        |  :|||  |  ::
Db        591 TCIVITVFIKHNNTPLVKASGRELCYILLFGVSLSYCMTFFFIAKPSPVICALRRLGLGT 650

Qy        643 PLTGCLSTLFLQA---AEIFVESELPLSWADRLSGCLRG--------PWAWLVVLLAM-L 690
                | | |  :    | ||           |  |          | : : : | : |
Db        651 SFAICYSALLTKTNCIARIF-------------DGVKNGAQRPKFISPSSQVFICLGLIL 697

Qy        691 VEVALCTWYLVAFPPEVVTDWHMLPTE---ALVHCRTRSWVSFGLAHATNATLAFLCFLG 747
```

```
Db        698 VQIVMVSVWLILETPG--TRRYTLPEKRETVILKCNVKD-SSMLISLTYDVVLVILCTVY 754

Qy        748 TFLVRSQPGCYNRARGLTFAMLAYFITWVSFVPLLANVQVVLRPAVQMGALLLC------ 801

Db        755 AFKTRKCPENFNEAKFIGFTMYTTCIIWLAFLPIF----YVTSSDYRVQTTTMCISVSLS 810

Qy        802 ---VLGILAAFHLPRCYLLMRQPGLN 824

Db        811 GFVVLGCLFA---PKVHIVLFQPQKN 833


RESULT 14
A42916
metabotropic glutamate receptor mGluR5 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A42916
R;Abe, T.; Sugihara, H.; Nawa, H.; Shigemoto, R.; Mizuno, N.; Nakanishi, S.
J. Biol. Chem. 267, 13361-13368, 1992
A;Title: Molecular characterization of a novel metabotropic glutamate receptor mGluR5 cc
A;Reference number: A42916; MUID:92317054; PMID:1320017
A;Accession: A42916
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1171 <ABE>
A;Cross-references: UNIPROT:P31424; UNIPARC:UPI0000000360; GB:D10891; NID:g220813; PIDN:
A;Experimental source: brain
A;Note: sequence extracted from NCBI backbone (NCBIN:107749, NCBIP:107750)
C;Keywords: G protein-coupled receptor; transmembrane protein

  Query Match             12.9%;  Score 584;  DB 2;  Length 1171;
  Best Local Similarity   24.3%;  Pred. No. 3.9e-37;
  Matches  218;  Conservative  154;  Mismatches  388;  Indels  136;  Gaps   37;

Qy         31 MKGDYVLGGLFPLGEAEEAGLRSRTRPSSPVCTRFSSNGLLWALAMKMAVEEINNKSDLL 90

Db         31 MPGDIIIGALFSVHHQPTVD-KVHERKCGAVREQY---GIQRVEAMLHTLERINSDPTLL 86

Qy         91 PGLRLGYDLFDTCSEPVVAMKPSLMFL---------AKAGSRDIAAYCNYTQYQPRVLAV 141

Db         87 PNITLGCEIRDSCWHSAVALEQSIEFIRDSLISSEEEEGLVRCVDGSSSFRSKKP-IVGV 145

Qy        142 IGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETFPSFFRTVPSDRVQLTAAAELL 201

Db        146 IGPGSSSVAIQVQNLLQLFNIPQIAYSATSMDLSDKTLFKYFMRVVPSDAQQARAMVDIV 205

Qy        202 QEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHEGLVPLPRADDSRLGKVQDVLH 261

Db        206 KRYNWTYVSAVHTEGNYGESGMEAFKDMSAKEGICIAHSYKI-YSNAGEQSFDKLLKKLR 264

Qy        262 QVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVASEAWLTSDLVMGLPGMAQMGTV 321

Db        265 S-HLPKARVVACFCEGMTVRGLL---MAMR---RLGLAGEF-----LLLGSDGWADRYDV 312

Qy        322 LGFLQRGA-----------QLHEFPQYV-----KTHLALATDPAFCSALGEREQ------ 359

Db        313 TDGYQREAVGGITIKLQSPDVKWFDDYYLKLRPETNL---RNPWFQEFWQHRFQCRLEGF 369

Qy        360 GLEEDVVGQRCPQCDCITLQNVSAGLNHHQTFS----VYAAVYSVAQALHNTLQCNASG- 414

Db        370 AQENSKYNKTCN--SSLTLR------THHVQDSKMGFVINAIYSMAYGLHNMQMSLCPGY 421

Qy        415 ---CPAQDPVKPWQLLENMYNLTF-HVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHD 470

Db        422 AGLCDAMKPIDGRKLLDSLMKTNFTGVSGDMILFDENGDSPGRYEIMNFKEMGK-----D 476

Qy        471 ------VGRF-NGSLRTERLKIRWHTSDNQKPVSR--CSRQCQEGQVRRV-KGFHSCCYD 520

Db        477 YFDYINVGSWDNGELKMDDDEV-WSKKNN---IIRSVCSEPCEKGQIKVIRKGEVSCCWT 532

Qy        521 CVDCEAGSYRQNPDDIACTFCGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLAL 580

Db        533 CTPCKENEYVF--DEYTCKACQLGSWPTDDLTGCDLIPVQYLRWGDPEPIAAVVFACLGL 590

Qy        581 GLVLAALGLFVHHRDSPLVQASGGPLACFGL---VCLGLVCLSVLLFPGQPSPARCLAQQ 637

Db        591 LATLFVTVIFIIYRDTPVVKSSSREL-CYIILAGICLGYLCTFCLI--AKPKQIYCYLQR 647

Qy        638 ---PLSHLPLTGCLSTLFLQAAEIFVESELPLSWADR--LSGCLRGPWAWLVVLLAMLVE 692

Db        648 IGIGLSPAMSYSALVTKTNRIARILAGSKKKICTKKPRFMSACAQLVIAFILICIQLGII 707

Qy        693 VALCTWYLVAFPPEVVTDWHMLPTEALVHCRTRSWVSFGLAHATNATLAFLCFLGTFLVR 752

Db        708 VAL----FIMEPPDIMHDYPSIREVYLICNTTNLGVVTPLGY--NGLLILSCTFYAFKTR 761

Qy        753 SQPGCYNRARGLTFAMLAYFITWVSFVPLL--ANVQVV-------LRPAVQMGALLLCVL 803

Db        762 NVPANFNEAKYIAFTMYTTCIIWLAFVPIYFGSNYKIITMCFSVSLSATVALGCMF---- 817

Qy        804 GILAAFHLPRCYLLMRQPGLNTPEFF---------LGGGPGDAQGQNDGNTGNQGK 850

Db        818 -------VPKVYIILAKPERNVRSAFTTSTVVRMHVGDGKSSSAASRSSSLVNLWK 866


RESULT 15
JC2132
metabotropic glutamate receptor 5 A - human
C;Species: Homo sapiens (man)
C;Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 05-Jan-1996
C;Accession: JC2132
R;Minakami, R.; Katsuki, F.; Yamamoto, T.; Nakamura, K.; Sugiyama, H.
Biochem. Biophys. Res. Commun. 199, 1136-1143, 1994
A;Title: Molecular cloning and the functional expression of two isoforms of human metabot
A;Reference number: JC2131; MUID:94197696; PMID:7908515
A;Accession: JC2132
A;Molecule type: mRNA
A;Residues: 1-1180 <MIN>
A;Cross-references: UNIPARC:UPI000017C2A2
C;Comment: This protein is coupled to guanine nucleotide binding proteins.
C;Keywords: glycoprotein; neurotransmitter; receptor; transmembrane protein
F;580-604/Domain: transmembrane #status predicted <TM1>
F;617-637/Domain: transmembrane #status predicted <TM2>
F;644-664/Domain: transmembrane #status predicted <TM3>
F;694-714/Domain: transmembrane #status predicted <TM4>
F;738-759/Domain: transmembrane #status predicted <TM5>
F;773-794/Domain: transmembrane #status predicted <TM6>
F;803-827/Domain: transmembrane #status predicted <TM7>

  Query Match             12.8%;  Score 580.5;  DB 2;  Length 1180;
  Best Local Similarity   24.5%;  Pred. No. 7.3e-37;
  Matches  219;  Conservative  151;  Mismatches  393;  Indels  131;  Gaps   37;

Qy         31 MKGDYVLGGLFPLGEAEEAGLRSRTRPSSPVCTRFSSNGLLWALAMKMAVEEINNKSDLL 90

Db         31 MPGDIIIGALFSVHHQPTVD-KVHERKCGAVREQY---GIQRVEAMLHTLERINSDPTLL 86

Qy         91 PGLRLGYDLFDTCSEPVVAMKPSLMF----LAKAGSRDIAAYC------NYTQYQPRVLA 140

Db         87 PNITLGCEIRDSCWHSAVALEQSIEFIRDSLISSEEEEGLVRCVDGSSSSFRSKKP-IVG 145

Qy        141 VIGPHSSELAMVTGKFFSFFLMPQVSYGASMELLSARETFPSFFRTVPSDRVQLTAAAEL 200

Db        146 VIGPGSSSVAIQVQNLLQLFNIPQIAYSATSMDLSDKTLFKYFMRVVPSDAQQARAMVDI 205

Qy        201 LQEFGWNWVAALGSDDEYGRQGLSIFSALAAARGICIAHEGLVPLPRADDSRLGKVQDVL 260

Db        206 VKRYNWTYVSAVHTEGNYGESGMEAFKDMSAKEGICIAHSYKI-YSNAGEQSFDKLLKKL 264

Qy        261 HQVNQSSVQVVLLFASVHAAHALFNYSISSRLSPKVWVASEAWLTSDLVMGLPGMAQMGT 320

Db        265 TS-HLPKARVVACFCEGMTVRGLL---MAMR---RLGLAGEF-----LLLGSDGWADRYD 312

Qy        321 VLGFLQRGA--------QLHEFPQYVKTHLALATD-----PAFCSALGEREQ------GL 361
```

```
Db        313 VTDGYQREAVGGITIKLQSPDVKWFDDYYLKLRPETNHRNPWFQEFWQHRFQCRLEGFPQ 372

Qy        362 EEDVVGQRCPQCDCITLQNVSAGLNHHQTFS----VYAAVYSVAQALHNTLQCNASG--- 414

Db        373 ENSKYNKTCN--SSLTLK------THHVQDSKMGFVINAIYSMAYGLHNMQMSLCPGYAG 424

Qy        415 -CPAQDPVKPWQLLENMYNLTF-HVGGLPLRFDSSGNVDMEYDLKLWVWQGSVPRLHD-- 470

Db        425 LCDAMKPIDGRKLLESLMKTNFTGVSGDTILFDENGDSPGRYEIMNFKEMGK-----DYF 479

Qy        471 ----VGRF-NGSLRTERLKIRWHTSDNQKPVSR--CSRQCQEGQVRRV-KGFHSCCYDCV 522

Db        480 DYINVGSWDNGELKMDDDEV-WSKKSN---IIRSVCSEPCEKGQIKVIRKGEVSCCWTCT 535

Qy        523 DCEAGSYRQNPDDIACTFCGQDEWSPERSTRCFRRRSRFLAWGEPAVLLLLLLLSLALGL 582

Db        536 PCKENEYVF--DEYTCKACQLGSWPTDDLTGCDLIPVQYLRWGDPEPIAAVVFACLGLLA 593

Qy        583 VLAALGLFVHHRDSPLVQASGGPLACFGL---VCLGLVCLSVLLFPGQPSPARCLAQQ-- 637

Db        594 TLFVTVVFIIYRDTPVVKSSSREL-CYIILAGICLGYLCTFCLI--AKPKQIYCYLQRIG 650

Qy        638 -PLSHLPLTGCLSTLFLQAAEIFVESELPLSWADR--LSGCLRGPWAWLVVLLAMLVEVA 694

Db        651 IGLSPAMSYSALVTKTNRIARILAGSKKKICTKKPRFMSACAQLVIAFILICIQLGIIVA 710

Qy        695 LCTWYLVAFPPEVVTDWHMLPTEALVHCRTRSWVSFGLAHATNATLAFLCFLGTFLVRSQ 754

Db        711 L----FIMEPPDIMHDYPSIREVYLICNTTNLGVVTPLGY--NGLLILSCTFYAFKTRNV 764

Qy        755 PGCYNRARGLTFAMLAYFITWVSFVPLL--ANVQVV-------LRPAVQMGALLLCVLGI 805

Db        765 PANFNEAKYIAFTMYTTCIIWLAFVPIYFGSNYKIITMCFSVSLSATVALGCMF------ 818

Qy        806 LAAFHLPRCYLLMRQPGLNTPEFF---------LGGGPGDAQGQNDGNTGNQGK 850

Db        819 -----VPKVYIILAKPERNVRSAFTTSTVVRMHVGDGKSSSAASRSSSLVNLWK 867
```

Search completed: December 16, 2005, 18:11:27
Job time : 50 secs

THIS PAGE BLANK (USPTO)